

# For Reference

---

NOT TO BE TAKEN FROM THIS ROOM

# For Reference

---

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS  
UNIVERSITATIS  
ALBERTAEENSIS



UNIVERSITY OF ALBERTA  
LIBRARY

Regulations Regarding Theses and Dissertations

Typescript copies of theses and dissertations for Master's and Doctor's degrees deposited in the University of Alberta Library, as the official Copy of the Faculty of Graduate Studies, may be consulted in the Reference Reading Room only.

A second copy is on deposit in the Department under whose supervision the work was done. Some Departments are willing to loan their copy to libraries, through the inter-library loan service of the University of Alberta Library.

These theses and dissertations are to be used only with due regard to the rights of the author. Written permission of the author and of the Department must be obtained through the University of Alberta Library when extended passages are copied. When permission has been granted, acknowledgement must appear in the published work.

This thesis or dissertation has been used in accordance with the above regulations by the persons listed below. The borrowing library is obligated to secure the signature of each user.

---









THE UNIVERSITY OF ALBERTA

COMPUTER ANALYSIS AND MAPPING OF SOIL TEST DATA

by



DOUGLAS RICHARD CAMERON, B.SC.

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE  
OF MASTER OF SCIENCE

DEPARTMENT OF SOIL SCIENCE

EDMONTON, ALBERTA

SPRING, 1969



UNIVERSITY OF ALBERTA  
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Computer analysis and mapping of soil test data" submitted by Douglas Richard Cameron, B.Sc., in partial fulfilment of the requirements for the degree of Master of Science.



## ABSTRACT

The Alberta Soil and Feed Testing Laboratory has received a total of 62,500 farm soil samples since April of 1962. The data concerning the farm soil samples were keypunched on computer cards and then transferred to magnetic computer tape. Techniques were devised and computer programs written to analyze the available information concerning the nutrient status of Alberta soils.

Summaries were made for nitrate nitrogen, phosphorus, potassium, and pH values for the top six inches of soil according to soil-climatic areas as well as for counties and other districts. The most accurate and promising summaries were the soil nutrient maps of Alberta showing areas of high and low nutrient concentration. The mapping of such data for the province involved trial runs with several different types of mapping techniques. The final selection was a circular weighted moving average technique, using townships as the basic mapping unit. The application of this technique resulted in the production of excellent and interesting soil nutrient maps displaying easily distinguishable trends. The same mapping function was applied to a smaller area, as well as a field, to determine further the nature of trends existing at these various levels. Careful consideration was given to crop types and time periods over which all summaries and mappings were conducted.

Preliminary work was carried out to determine the sampling accuracy associated with the soil test measurements, both on a field and regional basis. Phosphorus and nitrogen tests showed the most variability in sampling.





The A.S.F.T.L. data contained "ceiling values" for nitrogen and potassium which would strongly bias the results. This was moreso for potassium than for nitrogen. Models were developed and "probable" values were simulated to replace unknown values. The result was a much more plausible and accurate distribution of these data.

Large amounts of information were handled rapidly and efficiently with the aid of the electronic computer. All numerical calculations carried out within the manuscript were performed by programming the computer. Without the aid of this electronic device, only a small fraction of the results obtained and reported in this thesis would have been possible.



## ACKNOWLEDGEMENTS

Sincere appreciation is extended to Dr. J.A. Toogood, Head of the Department of Soil Science, for his guidance and encouragement in conducting this study, and for his valuable assistance and patience in the preparation and reviewing of the manuscript. Further thanks are extended to Dr. J.A. Robertson and Dr. M. Nyborg, staff members of the Department of Soil Science, and Dr. K.W. Smillie, staff member of the Department of Computing Science, for serving on the Committee and for assistance rendered.

Appreciation is also extended to the following: D.H. Lavery, Director of the Alberta Soil and Feed Testing Laboratory, for his helpful advice and guidance in the handling of the soil test data; J.A. Carson, Laboratory Supervisor, and the Laboratory technicians, A.S.F.T.L., for the special analyses they have carried out; G. Truscott, computer programmer, and the cardpunch operators, A.S.F.T.L., for their excellent processing of the data and other assistance kindly given; J. Gore, Economics Division, A.D.A., for his valuable work in sorting and merging the information on magnetic computer tape; A.W. Goettel, Head of the Soils Branch, Plant Industry Division, A.D.A., for his assistance in formulating the objectives of this study; and others who have rendered assistance.

A special thanks is extended to Mrs. Elaine Phillips for the typing of the manuscript. The author is deeply indebted to his fiancée, Miss Jacqueline Hirsch, for valuable aid in drafting and checking of the manuscript.



Acknowledgements are also extended to the following groups for their aid in the development of the thesis: the programming consultants employed by Information Services, Department of Computing Science; the computer operators and other staff of the Computing Centre; computing staff of the Department of Agricultural Economics, General Services Bldg.; Duplicating Services in the Henry-Marshall Tory Bldg.; Graduate Studies; and Printing Services.

The author wishes to record his appreciation to the Alberta Agricultural Research Trust and to those companies supporting the grant for the undertaking of this study.



# T A B L E O F C O N T E N T S

	Page
I. INTRODUCTION . . . . .	1
II. LITERATURE REVIEW. . . . .	3
Age of Computer Analysis . . . . .	3
Objectives of Soil Testing . . . . .	6
Soil Sampling . . . . .	7
Principles. . . . .	8
Laboratory vs. field variation. . . . .	9
Field size and number of samples. . . . .	10
Soil type and sampling . . . . .	13
Effects of time and fertilizer on sampling. . . . .	15
Effects of Crop, Fertilizer and Time on Nutrient Levels of the Soil . . . . .	19
Crop, fertilizer and time . . . . .	19
Summarization of Soil Test Data . . . . .	22
Mapping Techniques. . . . .	24
Regional and local trends . . . . .	24
"Eyeball" method. . . . .	26
Profile method. . . . .	26
Circle method. . . . .	27
"Expected value" method. . . . .	27
Abbreviated regression using orthogonal polynomials. . . . .	28
Complete regression using orthogonal polynomials. . . . .	29

(Continued)





	Page
Multiple regression method . . . . .	30
Simulation and Prediction of Unknowns. . . . .	30
Summary of Literature Review. . . . .	31
III. MATERIALS AND METHODS . . . . .	34
Materials. . . . .	34
Soil test data . . . . .	34
Cameron farm data. . . . .	40
Trial mapping data . . . . .	41
Other data. . . . .	41
Methods. . . . .	44
A. Data Processing. . . . .	44
B. Programming. . . . .	45
Sampling techniques. . . . .	45
Summarization techniques . . . . .	49
Mapping techniques. . . . .	49
Alberta mapping technique . . . . .	52
County and field mapping. . . . .	53
Revision and prediction of unknowns. . . . .	53
Other techniques. . . . .	54
IV. RESULTS AND DISCUSSION . . . . .	55
Soil Sampling . . . . .	55
Field basis . . . . .	55
Field size and sample numbers. . . . .	69
Sampling on a township and larger areas basis . . . . .	77

(Continued)



	Page
Summaries by Soil Area and County. . . . .	84
Preliminary. . . . .	84
Soil areas . . . . .	93
Counties. . . . .	129
Computer Mapping . . . . .	138
Results of the methods used. . . . .	138
Alberta soil nutrient maps . . . . .	149
County nutrient maps. . . . .	201
Field mapping. . . . .	206
Results of Prediction of Unknowns . . . . .	208
V.    SUMMARY AND CONCLUSIONS . . . . .	220
VI.   BIBLIOGRAPHY . . . . .	224
VII.  APPENDICES	



# LIST OF TABLES

Table		Page
I	Distribution of soil samples received by the A.S.F.T.L. based on date received . . . . .	34
II	Cumulative percentage distribution of number of cores taken to compose a farm soil sample . . . . .	56
III	Cumulative percentage distribution of field size . . .	57
IV	Cumulative percentage distribution of acres per core .	57
V	Percent accuracy of nitrogen sampling on the Cameron farm data as determined by simulation and statistical techniques. . . . .	60
VI	Percent accuracy of phosphorus sampling on the Cameron farm data as determined by simulation and statistical techniques. . . . .	61
VII	Percent accuracy of potassium sampling on the Cameron farm data as determined by simulation and statistical techniques. . . . .	63
VIII	Percent accuracy of pH sampling on the Cameron farm data as determined by simulation and statistical techniques. . . . .	64
IX	Average number of cores required for nitrogen sampling on different field sizes . . . . .	70
X	Average number of cores required for phosphorus sampling on different field sizes. . . . .	71
XI	Average number of cores required for potassium sampling on different field sizes. . . . .	72
XII	Average number of cores required for pH sampling on different field sizes. . . . .	73
XIII	Average number of samples required within a township and critical circle for nitrogen, phosphorus, potassium, and pH on different crop groupings and years for 80% sampling accuracy. . . . .	78
XIV	Average number of samples within a township and critical circle for different crop groups and months and years ..	79

(Continued)



Table		Page
XV	Summary of the nutrient availability of Alberta soils under fallow and stubble crops over different years. . .	85
XVI	Percentage distribution of crops by area. . . . .	87
XVII	Distribution of soil samples received by the A.S.F.T.L. based on year sampled. . . . .	89
XVIII	Distribution of soil samples received by the A.S.F.T.L. based on the month sampled. . . . .	89
XIX	Results of yearly variation of some selected treatments on the economic productivity sites. . . . .	92
XX	Percentage distribution of potassium "ceiling values" by soil area for the A.S.F.T.L. data collected in time period of April, 1965 to March, 1969. . . . .	211
XXI	Percentage distribution of potassium values above 600 lb/ac by soil area. . . . .	214
XXII	Percentage distribution of potassium values above 800 lb/ac by soil area. . . . .	214
XXIII	Percentage distribution of potassium values above 1000 lb/ac by soil area. . . . .	214
XXIV	Percentage distribution of potassium values between 601 and 800 lb/ac, inclusive. . . . .	215
XXV	Percentage distribution of potassium values between 801 and 1000 lb/ac, inclusive . . . . .	215





Figure		Page
14.	Comparison of phosphorus averages and standard deviations on different crop types for Area 9 . . . . .	107
15.	Phosphorus averages and standard deviations by crop grouping and soil area. . . . .	110
16.	Yearly variation of phosphorus on stubble and fallow in Area 6 . . . . .	111
17.	Phosphorus averages on stubble by soil area and sampling year . . . . .	113
18.	Comparison of potassium averages and standard deviations on different crop types for Area 2 . . . . .	116
19.	Comparison of potassium averages and standard deviations on different crop types for Area 6 . . . . .	117
20.	Comparison of potassium averages and standard deviations on different crop types for Area 9 . . . . .	118
21.	Potassium averages and standard deviations by crop grouping and soil area . . . . .	119
22.	Yearly variation of potassium on stubble and fallow in Area 6 . . . . .	121
23.	Potassium averages on stubble by soil area and sampling year. . . . .	122
24.	Comparison of pH averages and standard deviations on different crop types for Areas 2, 6, and 9 . . . . .	124
25.	Comparison of pH averages and standard deviations by crop grouping and soil area . . . . .	125
26.	Yearly variation of pH on fallow and stubble in Area 6 . . . . .	127
27.	Averages of pH on stubble by soil area and sampling year . . . . .	128
28.	Map of Alberta showing the sampling density in counties and districts . . . . .	131
29.	Map of Alberta showing nitrate nitrogen averages on fallow in counties and districts . . . . .	132
30.	Map of Alberta showing nitrate nitrogen averages on stubble in counties and districts . . . . .	133

(Continued)



Figure		Page
31.	Map of Alberta showing phosphorus averages on stubble in counties and districts . . . . .	135
32.	Map of Alberta showing potassium averages on stubble in counties and districts . . . . .	136
33.	Map of Alberta showing pH averages on stubble in counties and districts. . . . .	137
34.	Map of trial area displaying original values of phosphorus rounded-off and contoured . . . . .	139
35.	Map of trial area displaying results of the profile method . . . . .	141
36.	Map of trial area displaying results of the circle method . . . . .	143
37.	Map of the trial area displaying results of the "expected value" method . . . . .	145
38.	Map of the trial area displaying the results of the complete regression method using orthogonal polynomials . . . . .	147
39.	Map of Alberta showing the average number of samples per township . . . . .	154
40.	Map of Alberta showing distribution of sample numbers on all crops for the time period . . . . .	155
41.	Map of Alberta showing nitrate nitrogen values on fallow for townships with 2 or more samples available. . . . .	156
42.	Map of Alberta showing nitrate nitrogen values on stubble and grass-legume crops for townships with 2 or more samples available . . . . .	157
43.	Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: all months, 1962-69. . . . .	160
44.	Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: all months, 1962-69. . . . .	161
45.	Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1968 - March, 1969 . . . . .	163

(Continued)



Figure		Page
46.	Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: Sept., 1968-March, 1969 . . . . .	164
47.	Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1967 - June, 1968 . . . . .	165
48.	Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: Sept., 1967- Apr., 1968 . . . . .	166
49.	Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops, using larger rounding factors, for the time period: July, 1967 - June, 1968 . . . . .	167
50.	Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1966 - June, 1967 . . . . .	168
51.	Map of Alberta showing distribution of nitrate nitrogen on stubble and legume grass crops for the time period: Sept., 1966-Apr.,1967 . . . . .	169
52.	Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1965 - June, 1966. . . . .	170
53.	Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: Sept., 1965-Apr., 1966 . . . . .	171
54.	Map of Alberta showing phosphorus values on fallow, stubble, and grass-legume crops for townships with 2 or more samples available . . . . .	174
55.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: Apr., 1962-Mar.,1969 . . . . .	175
56.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops, using larger rounding factors, for the time period: Apr., 1962 - Mar., 1969 . . . . .	176
57.	Map of Alberta showing distribution of phosphorus on all crops, using a smaller circle, for the time period: Apr., 1962 - Mar., 1969.	178

(Continued)





Figure		Page
58.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1968 - Mar., 1969 . . . . .	179
59.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1967 - Mar., 1968 . . . . .	180
60.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1966 - June, 1967 . . . . .	181
61.	Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1965 - June, 1966 . . . . .	182
62.	Map of Alberta showing potassium values on all crops for townships with 2 or more samples available. . . . .	184
63.	Map of Alberta showing distribution of potassium on all crops for the time period: Apr., 1965 - Mar., 1969 . . . . .	185
64.	Map of Alberta showing distribution of potassium on all crops, using a smaller circle, for the time period: Apr., 1965 - Mar., 1969 . . . . .	187
65.	Map of Alberta showing distribution of potassium on all crops for the time period: July, 1968 - Mar., 1969. . . . .	188
66.	Map of Alberta showing distribution of potassium on all crops for the time period: July, 1967 - June, 1968. . . . .	189
67.	Map of Alberta showing distribution of potassium on all crops for the time period: July, 1966 - June, 1967. . . . .	190
68.	Map of Alberta showing distribution of potassium on all crops for the time period: July, 1965 - June, 1966. . . . .	191
69.	Map of Alberta showing pH values on all crops for townships with 2 or more samples available . . . . .	193
70.	Map of Alberta showing distribution of pH on all crops for the time period: Apr., 1962 - Mar., 1969. . . . .	194

(Continued)





Figure		Page
71.	Map of Alberta showing distribution of pH on all crops, using a smaller circle, for the time period: Apr., 1962 - Mar., 1969 . . . . .	196
72.	Map of Alberta showing distribution of pH on all crops, for the time period: July, 1968 - June, 1969 . . . . .	197
73.	Map of Alberta showing distribution of pH on all crops, for the time period: July, 1967 - June, 1968 . . . . .	198
74.	Map of Alberta showing distribution of pH on all crops, for the time period: July, 1966 - June, 1967 . . . . .	199
75.	Map of Alberta showing distribution of pH on all crops, for the time period: July, 1965 - June, 1966 . . . . .	200
76.	Map of the County of Camrose showing isopleths of phosphorus content on fallow, stubble, and legume-grass crops. . . . .	202
77.	Map of the County of Camrose showing isopleths of potassium content on all crops . . . . .	204
78.	Map of the County of Camrose showing isopleths of soil pH on all crops . . . . .	205
79.	Nutrient maps of the Cameron farm field showing distribution of nitrate nitrogen, phosphorus, potassium, and pH . . . . .	207
80.	Observed and theorized distributions of high nitrate nitrogen values previously recorded as 80 +. . . . .	209
81.	Diagramatic flowchart of the computer simulation of unknown potassium values . . . . .	213
82.	Potassium percentage distribution by soil area. . . . .	217
83.	Map of Alberta showing distribution of phosphorus where both samples and townships have been weighted . . . .	242
84.	Map of Alberta showing distribution of phosphorus where averages per township have been calculated before applying the mapping function. . . . .	243



## L I S T O F A P P E N D I C E S

Appendix		Page
I	Format of A.S.F.T.L. data . . . . .	230
IIa	Diagram and description of the profile mapping technique. . . . .	235
IIb	Diagram and description of the circular mapping technique. . . . .	237
III	Description of the Alberta mapping program . . .	240



## I. INTRODUCTION

There is a real need for obtaining comprehensive and reliable data on the nutrient status of soils of any country. Such information would be helpful to both the farmer and the agricultural leader in revealing the general nutrient status of the soils in their community. A general picture of the nutrient status of the soils would aid research workers in orientating their research and possibly open up new areas of research. The data would give educational and action agencies, as well as fertilizer industries, a sound basis for evaluating or promoting their respective programs. Furthermore, if such data were obtained at appropriate intervals, they would reflect trends in soil fertility. The data would give information as to what soil areas are being improved and which are on the decline. Trends in nutrient status would reflect the effectiveness of research, educational, and action programs. The data would also serve as a broad basis for estimating fertilizer and lime needs for given areas.

Since the Agriculture Soil and Feed Testing Laboratory was established in 1956, over 75,000 soil samples have been collected and analyzed. Much information can be obtained from these data, but summarization is tedious and time consuming even when using the more advanced types of desk calculators. The development of the modern digital computer simplifies the task immensely. The application of computer techniques to a study of Alberta soil test data is the subject of this thesis.

The analyses of the soil test data include summaries and preparation of maps of Alberta and areas within Alberta showing levels of nitrogen, phosphorus, potassium and pH in the top 6 inches of Alberta soils. Sampling methods, cultural practices, and time of collecting samples are very



carefully considered in this study.

The project was undertaken with the following objectives:

1. To determine which samples can be considered to be reliable, based on various studies of sample numbers required to give a reasonable estimation of the true mean,
2. To develop programs that can be used to summarize the data according to soil-climatic zones, and districts as well as cropping practices and seasons of the year,
3. To determine the most applicable and accurate mapping technique for the various kinds of data available in Alberta's Soil Testing program and apply this technique to the data, and
4. As a consequence of the above objectives, to assist the Laboratory in its operations.

This study should make it possible for those who interpret soil test data to do a better job and make possible more accurate and practical recommendations for fertilizer use.





## II. LITERATURE REVIEW

### Age of Computer Analysis

Since the dawn of history man has been processing data. Within the last 50 years the amount of data has grown to such proportions that it has swamped man's ability to handle it (Thomas, 1965). Within the last decade a "thinking revolution" and an "information explosion" have occurred. The battle of coping with the masses of data has mainly been won because man has devised a machine to relieve him of mental drudgery - the electronic computer.

Computers play a key role in the revelation of new facts and in their ability to store information for later recall. Although computers have been used in various aspects of agriculture for the past decade, it has not been until just recently that this use has begun to intensify. Favret (1965) points out that the most important aspect of the use of computers in agriculture thus far has been in the field of statistics, but many uses are rapidly emerging involving the application of agricultural and economic data, especially along the line of linear programming and simulation of events using models or "Monte Carlo" techniques.

Shickluna and Walsh (1967) and Walsh (1964) discuss the use of electronic data processing equipment in the Michigan and Wisconsin soil testing programs. In those states accurate fertilizer recommendations are made by computer, based on laboratory test results and such factors as soil type, soil management group, yield potential, subsoil fertility, cropping sequence, crop variety, subsoil moisture, weather probabilities,



and the managerial ability of the farmer. Electronic data processing equipment can, obviously, serve as a valuable agronomic tool for interpreting soil test results and for estimating the amount of lime and fertilizer to be applied. By integrating all available data on soil fertility and plant nutrition relationships with soil test results more reliable interpretations should be possible.

A by-product of computer-handling of recommendations is that the soil test information is in a form which can be utilized later for preparation of up-to-date summaries. The use of high speed computers makes it possible to study the inter-relationships among the various physical and chemical tests that are carried out on routine samples submitted to the soil testing laboratory. Computers make possible a new approach for studying the broad spectrum of nutrient balance. In this approach analytical results from large numbers of soil samples can be integrated into various specific research programs. Many new avenues for investigations will undoubtedly appear as computer techniques are refined.

Van Bavel (1966) gives a very excellent report on the use and abuse of information processed by machine, and it is highly recommended reading for all researchers working with computers. He states that when digits are faithfully clicking and tapes are humming, printing out columns of neat figures (and papers must be published) there can certainly be a great impetus for research with little theoretical background. This is further encouraged by statistical analyses that are available in "canned" form. Statistical tools are important but it is pertinent that many research biologists, agriculturalists, and hydrologists have been



brought up in a pattern of thinking that a standard statistical treatment can redeem one's inability to imagine critical questions that penetrate the complexity of biological processes or geophysical phenomena.

Van Bavel emphasizes the point that the quality of the computer output is no better than that of the input. Formats and programs in computing should allow flexibility and access to primary data - implying the utmost simplicity compatible with adequate evaluation of data. Since programming is generally done by a specialist who may not have an appreciation of the research problem, close cooperation between the investigator and programmer is necessary during writing and "debugging". Research applications are not likely to be based on continuous use of the same program and it is expected that the researcher will change or scrap the program as his work evolves; the less detailed the program, the fewer changes that will need to be made. Machine processing of data should not make scientists busier but more effective. Lastly, final assessment of the significance of any body of data will always involve human judgement, a factor that no automatic device can presently replace.

Over the past 10 years the speed at which computers calculate has increased by a factor of a million. Tomorrow's machines will no doubt be doing the same things even faster and venturing into new areas of use. The future for computer applications in agriculture is almost unlimited. Thomas (op. cit.) states that these electronic wizards will be involved in food production - right from the initial phases of agricultural research, through production, to processing and distribution. It is even envisioned that many farm operations will be entirely under electronic control.





### Objectives of Soil Testing

Melsted (1967), speaking in general terms, states that soil testing is any chemical or physical measurement made on a soil. But through common usage the term "soil testing" has been given both a more restricted and a much broader meaning - restricted in the sense that it has come to mean rapid chemical analyses to assess the available nutrient status of a soil, and broadened to include interpretations, evaluations, and fertilizer recommendations based on the chemical analyses and other considerations.

One of the first recorded works of soil testing was that of Liebig in 1850 (Peterson and Calvin, 1965). From that time up to the early 1920's little progress was made until significant contributions were made to soil chemistry and worthwhile soil tests developed by such scientists as Dyer, Hilgard, Burd, Bray, Hester, Morgan, Spurway, and Truog (Melsted, op. cit.). Since the late 1940's soil testing has been widely accepted as an essential tool in formulating a sound fertilizer program. The Report of the ARDA Soil Test Appraisal as conducted by Toogood and Anderson (1964) in June and July 1963 in Alberta certainly bolsters the above statement. The report shows that 61% of the farmers surveyed carefully followed the recommendations made by the Testing Laboratory. The main reason for not doing so was limited capital. Despite the fact that 50% suffered from below average growing conditions, 68% found that fertilization was profitable in the year tested. As for continuing with soil testing, 91% of the farmers indicated that they planned to do so.

Soil testing is not universally accepted however. Dechering (1956)





points out that in almost exclusively agricultural countries, such as Pakistan, soil testing will produce the least effect, due to the fact that no facilities or funds are available to follow up recommendations made and, adding to the problem, the level of overall farm production and education is much lower than in industrialized countries. In the latter soil testing is more widely accepted and utilized.

A soil testing program may be divided into four phases: collecting the soil samples, extracting and determining the available nutrients, interpreting the analytical results, and making fertilizer recommendations. The objectives or aims of soil testing, then, according to Melsted (op. cit.) are:

1. To determine accurately the available nutrient status of soils,
2. To indicate clearly to the farmer the seriousness of any deficiency, particularly with respect to various crops,
3. To form the basis on which fertilizer needs are determined, and
4. To express the results in such a way that they permit an economic evaluation of the suggested fertilizer recommendations.

### Soil Sampling

A soil testing program starts with the collection of a soil sample or samples from a field. This is the critical part, for regardless of the care taken in the analysis, and in spite of the large amount of data that may be collected to establish benchmarks and to provide chemical tests for availability indices, the soil test will be of little value if the sample submitted by the farmer is not representative of the area to



which the recommendation is to be applied (Henry, 1967). Similarly any summary of such data, whether it be tabular, statistical or mapping, will only be as accurate as the samples it represents. This very fact has prompted a review of the literature dealing with soil sampling.

Principles: Cline (1944) wrote a classical paper dealing with the principles of soil sampling. He states that the accuracy to which a soil sample represents the population sampled depends upon soil variability, the number of sampling units contributing, and the way in which the sample is drawn.

Three general principles should be kept in mind when selecting sampling units:

- (a) A sample composed of a few sampling units scattered at random throughout a homogeneous population contains information up to the limits of its size, but even a large sample, confined to a part of the population, contains no information about excluded parts.
- (b) An unbiased estimate of the mean requires that every sampling unit have an equal chance of being drawn.
- (c) An unbiased estimate of significance and fiducial limits requires that every sample of  $n$  sampling units have an equal chance of being drawn.

Composite samples are valid only if

- (a) The sampling volume represents a homogeneous population,



- (b) Equal amounts of each sampling unit contribute to the sample analyzed,
- (c) No interactions that would affect the results materially occur, and
- (d) An unbiased estimate of the mean is the only objective.

Similar principles to those that Cline mentions are found in part throughout the literature available on soil sampling. Kempthorne and Allmaras (1965), Allmaras (op. cit.), Dixon (1965), Black (1965), and Petersen and Calvin (op. cit.) cover the principles of soil sampling and associated bias and variation even more thoroughly than Cline. Although not necessarily dealing with soils, other workers such as Cochran (1953), Sukhatme (1954), Snedecor (1956), Steele and Torie (1960), and Sampford (1962) have contributed significantly to sampling theory. These authors have been frequently referred to by soil scientists, mainly for the statistical techniques that they have presented in their books.

Laboratory vs. field variation: Cline (op. cit.) states that the limit of accuracy generally is determined by the sample, not by the analysis. From his experience, probable errors for sampling are 3 to 6 times greater for sampling and sample treatment than for subsampling and analysis. Reed and Rigney (1947) have shown it to be several times greater than analytical error for a variety of soil properties. Pizer (Hemingway, 1955) has suggested that improper soil sampling is of frequent occurrence and is the main factor to be checked when analytical results and plant symptoms do



not agree. Ferrari and Vermeulen (1956) of the Netherlands point out that the error of analysis, generally speaking, is small and does not play a great part when compared with other errors. An error of 4% can be expected. Hemingway (op. cit.) shows that laboratory variations are small compared to field variation. He also points out that there are definite variations in sampling error, itself, depending upon the nature of the nutrient being measured, the mean of the determination, and the application of fertilizers or manure to the field being sampled.

The unfortunate part of the work done by these workers is that the term "sampling error" is not clearly defined and should theoretically not be used as a comparison measurement since it varies with the number of sampling units or cores taken from a field. In addition, there is the possibility that it might be affected also by field size.

Field size and number of samples: The older literature dealing with the number of individual cores that should be taken from a field is not only diverse, but often fails to indicate the statistical reliability of the method employed. Hemingway (op. cit.) states that both Pizer and Hunt, Swanson and Jacobson regard 10-20 sampling units as sufficient to form a reliable sample. Cline (op. cit.), however, gives 20 as a minimum. According to Ferrari and Vermeulen (op. cit.) the instructions in the Netherlands prescribe 30-40 borings. In this case they claim the sampling error is reduced by about 80%.

Riehm (1956) suggests that in practice 20 individual samples be taken from an area of 3 hectares (7.4 acres) for the mixed sample and 15-20





individual samples from an area of 1 hectare (2.47 acres), while 10-15 individual samples are sufficient for an area of 1/4 hectare (.62 acre).

Reed and Rigney (op. cit.) indicate that 30 borings in a visually uniform field of 5 acres would permit the determination of phosphorus, calcium, and potassium to within 10% of the true field average.

Hemingway (op. cit.) reports that the sampling error does not increase as the field size increases. He found that in almost every case 24 sampling units were needed to give a reliable mean for potassium, and fewer than 10 if an error in the mean pH of  $\pm 0.5$  is permissible. Tisdale and Nelson (1966) suggest that the composite sample which is equivalent to a pound from an average of 10 acres (1:20,000,000 lbs. of surface soil) is extremely small and would likely cause considerable error. They, however, do recommend that the composite need only consist of 15-20 borings from 10 acres. The Alberta Soil and Feed Testing Laboratory suggests 15-20 sampling sites for a composite sample from a field no larger than 60-70 acres.

Discussions of sampling programs based on statistical analysis of soil fertility data can be found in recent literature. The following reviews which will be discussed briefly, are important for their statistical techniques used to obtain an estimation of the sample numbers required, rather than for the application of their results to other data.

Rigney and Reed (1946) studied the variation of 7 chemical properties in 20 adjoining fields and discussed the effect of varying the number of samples and subsamples on the standard error of the estimated mean. Reed and Rigney (op. cit.) give the number of samples and subsamples necessary to achieve specified limits of precision for various properties and report



that the properties studied have different patterns of variation. Jacob and Klute (1956) give variance components for plots, samples, and determinations for several chemical and physical properties of Sassafras sandy loam on Long Island. They conclude that because of the varying ratios between variance components for different characteristics, different optimum sampling plans can be found. The current status of soil sampling techniques was reviewed by Rigney (1956) and nicely sums up the work that had been carried out in soil sampling up to that time.

The above authors state in general that the accuracy of a prediction may be represented by the following equation:

$$\bar{V_x} = \frac{V_p}{p} + \frac{V_b}{b} + \frac{V_s}{s} + \frac{V_a}{a}$$

Here the variance ( $\bar{V_x}$ ) of a mean value for a particular area is composed of the sum of the variances arising from the different steps in the sampling process:  $V_p$ ,  $V_b$ ,  $V_s$ , and  $V_a$  are estimates of the true variances obtained for positions, borings, subsamples, and aliquots.

Hammond, Pritchett and Chew (1958) made a study of some of the chemical and physical properties of 7 fields varying in size from 2 to 35 acres using the multistage sampling design. Two and three-stage samplings were compared to simple random sampling, and optimum (i.e. minimum cost) allocation of samples were determined for estimates of certain soil properties with given limits of precision. They have coined the words "macrouniformity" and "microuniformity" as related to soil heterogeneity. The degree of "macrouniformity" is shown by the efficiency of multistage



relative to simple random sampling and fields with greater than 50% efficiency were generally considered "macrouniform".

Mader (1963) discusses the variability in soils sampled in a red pine plantation site study. His results show that some soil properties require a large number of samples for accurate estimation, others a small number. Leo (1963) studied the heterogeneity of soil fertility in apparently uniform agricultural land. The study reveals that the number of sampling units we are currently recommending for soil testing is far too low. It is the opinion of the author that the significance levels and precision levels used in this study were too restrictive, thus giving exceedingly high results with respect to number of sampling units required.

Recently Holland, Little, Allen and Dermott (1967) report that about 24 cores should be taken at random throughout an orchard in such a way that the rows and alleyways contribute equally. The precision of the resulting determinations will be such that the true value appropriate to an orchard will, in the long run, lie within  $\pm 15\%$  of the observed value on 4 occasions out of 5 (90% confidence limits), with the exception of pH, which is more precisely determined ( $\pm 4\%$ ) and soil phosphorus which is less precisely determined ( $\pm 25\%$ ).

Soil type and sampling: Cline (op. cit.) states that soil bodies should be subdivided horizontally to provide sampling areas that are homogeneous with respect to soil type, plant growth, and treatment. In addition, the recognizable horizons of the same soil type represent different populations both physically and chemically and are logical units of vertical





subdivision. Many of the conflicting results in the literature may partially be due to the mixing of unlike horizons. Welch and Fitts (1956) carried out a study on the effect of depth on sampling. Although they did not sample by horizon, their work indicates strongly that different horizons affect the nutrient level of the sample. Their work shows a significant decrease in the pH of the soil with depth from the 0 to 3 to the 3 to 6 to the 6 to 9 inch layers of soil in fields seeded to sod crops. However, their data show no significant difference for comparable fields used for row crops. The data for phosphorus, potassium, and organic matter are significantly lower for samples composited from 0 to 6 inch cores as compared to 0 to 3 inch cores from fields in sod crops as well as fields in row crops.

Ferrari and Vermeulen (op. cit.) state that big differences are sometimes found on resampling and retesting a field either with or without an interval of time between the samplings. These differences may therefore, be largely due to the fact that the soil in the field is heterogeneous. The authors also imply that the differences in homogeneity between the various types of soil are usually small. Their results do, however, point out that loess and alluvial soils are less homogeneous in some respects than are other types of soils.

Peck and Melsted (1967) show from the work of Lavery in central Illinois on detailed field sampling, that except for one general agreement of alkaline pH's associated with a particular soil type, there is no indication of pH or phosphorus values being related to soil type within a field. Contrary to this, Henry (op. cit.) indicates from his work in





Saskatchewan that soil sub-groups have a more or less definite available nutrient status. His work indicates that Gleysolic soils have a fairly high phosphorus status but respond to added phosphorus, and therefore, should be neglected when establishing a sampling pattern. Parker et al. (1951) in summarizing the soil test values of Tennessee found a relationship between the soil association map of Tennessee and the available phosphorus in the soil. Their summary also shows relationships between soil associations (and their parent materials) and levels of potassium and pH where the effect is not marred by fertilization and cropping practices.

Soil type, itself, is quite variable both morphologically and chemically. In many instances a field is composed of quite a few soil types and even if these were tested separately, the fertilization of them separately would be impractical to the ordinary farmer.

Effects of time and fertilizer on sampling: Tisdale and Nelson (op. cit.) state that since levels of lime, phosphorus, and potassium in the soil normally do not change rapidly, samples for lime and fertilizer recommendations do not have to be taken just before the crop is planted. Instead, samples may be taken in the late summer or fall when there will be ample time to obtain the needed fertilizer for spring-sown crops.

Riehm (op. cit.) in Germany found when sampling the same fields over time that no significant fluctuations in pH and nutrient content occur in untilled and unmanured ground. Even after several years the pH, phosphate, and potassium values were almost equally high. However, he states that fluctuations on cropland were greater after mineral fertilizer had been applied. On manured ground he advises that sampling should only



be carried out one or two months after manuring. In view of this, sampling on unmanured ground is justified the whole year round and this is only slightly less true for cultivated ground.

Horton and Stinson (1939) in working with tobacco where heavy fertilization is carried out in or near the row found that at least 100 sampling units must be used 3 weeks after fertilization to obtain a representative sample. Ten weeks after fertilization the number of places may be reduced to 50. This number may be reduced to 25 one or more years after fertilization. The results of Hemmingway (op. cit.) also point out the effects of fertilizers on sampling. He states that soils which have received lime and fertilizers within 3 years of sampling showed appreciable greater sampling errors than those which had not.

Bell and Thornton (1937) in testing soils over time with different applications of fertilizer found no consistent change in soil pH with season nor any change due to fertilization. Nitrate nitrogen was increased by fertilization and fluctuated considerably during the cropping season, reaching a maximum during July. Phosphorus and potassium supplies in the soil were increased significantly by fertilizer treatments and both of these nutrients showed a tendency to decrease as the season advanced. This was particularly evident with potassium. When the summer was hot and dry, conditions apparently were conducive to an extreme decline in the amount of available potassium in the soil.

Raupach (1951a, 1951b) working in Australia found that there was no reason to believe that the seasonal variation of pH, soluble salts and other factors will be the same from place to place. Since some of the



factors governing a variation with season may also contribute towards a variation in space, the picture of a complex pattern of pH gradients continually varying with time in a non-uniform manner is not unreasonable. Raupach (1951a) could not find any correlation of rainfall data with pH or total soluble salts. On the other hand, Carolus and Lucas (1943) point out that there is a definite drop in pH from spring to summer and then an increase from summer to fall. Their results also show that rainfall and leaching increased the pH of dry soils. Swanback and Morgan (1930) working in Connecticut found that during the late spring and early summer season, particularly after high applications of nitrogenous fertilizers, there is a marked increase in soil acidity (decrease in pH). A dry season produces a more acid condition than a wet season. They conclude that the best time to test soils for pH would be in April or October since intermediate values are attained then.

Nitrate nitrogen is probably the most variable nutrient with time. Filingier (1931) attributed the conflicting results reported for the behavior of nitrogen in orchard soils mainly to the lack of uniformity in time of sampling. There are a number of factors that may be responsible for the conflicting results obtained. First, the type of soil no doubt has an influence on the behavior of nitrates throughout the season. The type of soil influences the water content, the nature of the organic matter, and the organisms in the soil and thus the nitrate content. Second, the weather conditions influence the nitrate accumulation. Heavy rains carry the soluble nitrates to lower soil strata. Low temperatures retard, whereas, high temperatures accelerate nitrate formation. As work is done in different parts of the country, different





soil types and weather conditions no doubt will explain part of the conflicting results. A third factor, influencing the results obtained by investigations, is the frequency of taking soil samples and the depth to which they were taken. Filinger recommended that an investigator should take soil samples at regular short intervals of not more than a week and to a depth of 2 to 3 feet to get a true picture of the behavior of nitrates in an orchard soil.

In the 1920's and 1930's much work was carried out in the nutrient supplying capacity of Alberta soils especially with nitrate supplying power under various crops and throughout the seasons. Wyatt, Ward and Newton (1926) state that there is a distinct relationship between the moisture content and nitrate content of the soil, especially during the early part of the summer and until the rapid rate of plant growth disturbs this relationship. In the case of summerfallow this relationship persists into the summer and becomes disturbed only when the soil moisture becomes sufficiently great to leach some of the nitrates into lower layers of soil. However, in general, the nitrate production shows a slight lag behind the moisture and temperature curves. There is a major maximum in nitrate production during the later part of spring and the early part of summer, and a minor maximum during the early part of the fall. However, when crops are occupying the soil the high point in nitrate accumulation seldom extends into the summer season. Newton (1954) showed that soil nitrate does not accumulate appreciably in the spring until the surface soil temperature reaches about 15° C, but ammonia does frequently accumulate in black soil below this temperature. Accumulation of ammonia nitrogen has been measured also in the late fall, before freeze-up.





Effects of Crop, Fertilizer and Time on Nutrient Levels  
of the Soil

A review of the effects of cropping practices, fertilizer and time on the chemical test values of the soil is important for three main reasons:

1. The samples received by soil testing laboratories come from soil under different cropping practices, sampled in various months over a period of years.
2. Summarizations of such data are usually carried out for the different nutrients with respect to particular crops or crop groupings within certain time periods. Such summaries promote the understanding of the crop, time, and fertilizer inter-relations.
3. To help decide which crops or crop groupings and which time periods one should choose with respect to the particular test value being summarized or mapped.

Crop, fertilizer and time: Tisdale and Nelson (op. cit.) using data compiled by Welch and Kamprath of North Carolina show that for garden crops, tobacco, and cotton, soil phosphorus levels are usually high. The reason for this they claim, is that these crops receive 50 to 60 lb/ac of phosphorus annually. Soils producing small grains show about medium levels of phosphorus, while clovers and grasses show relatively low levels of phosphorus. Peevy, Smith, and Brown (1940) of Iowa, in summarizing their



rotational treatment plots over 21 years, showed that in general, losses of total phosphorus were insignificant, however, soils treated with rock phosphate showed highly significant gains in phosphorus indicating an accumulation of phosphorus in the surface soil from the fertilization. Parker et al. (op. cit.) in summarizing the soil test data of North Carolina and Tennessee over counties have noted a high phosphorus index following tobacco and cotton crops but a relatively low index following pastures.

There has been little work done on the effect of cropping and fertilization on soil tests for potassium. Farrari and Vermeulen (op. cit.) state that the average potassium content of grassland is higher than that of arable land in the Netherlands. Parker et al. noted that there was little relation between the potassium indices and crops. These authors also showed that a study by Mehring and Parks revealed that fertilizers supplied only 52% as much  $K_2O$  as was removed by crops. Hence, as compared to phosphorus, the residual effect of the fertilizers on the general potassium levels of soils would be expected to be relatively small and they concluded that on a broad basis, the content of potassium in the soil is more closely related to soil association than it is to fertilization. Nelson (1964), in working with Alberta soils with low potassium, found that continuous cropping in the greenhouse did not reduce the potassium in the majority of soils used. His work also showed that the addition of potassium at 60 lb/ac of K before each crop increased the potassium content of each soil.

Little is known of the effects of crop type on pH. The work of Parker et al. indicates that acidity is well correlated with liming



practices in both Tennessee and North Carolina. Work by Toogood, Bentley, Webster and Moore (1962) indicates that high yielding 5-year rotational plots have lower pH's than low yielding plots. In addition, pH is lower, on the average, in plots cropped to the 5-year rotation than in the wheat fallow series. Swanback and Morgan (op. cit.) state that high applications of nitrogenous fertilizers tend to decrease pH.

The work of Wyatt et al. (1926, 1927) and Newton (1954) gives an excellent summary of the effect of cropping practices and fertilizers in soils of Alberta. The result of their investigations showed that the production of nitrates in field soils is influenced by the crop growing, the crop sequence, and the method of tillage, together with moisture and temperature factors. The perennial crops such as timothy and alfalfa keep nitrates at a lower level than do the annual, such as wheat, barley, corn and potatoes. The non-tilled annuals such as wheat and barley keep the nitrates slightly lower than the intertilled crops such as corn and potatoes. Summerfallow showed much greater accumulations of nitrates than were found under corn or potatoes. The plant residues from legumes promoted nitrification more vigorously and at an earlier date than did residues from non-legumes. In the black earth loam, for example, nitrates were reduced to low levels under alfalfa and grasses, indicating relatively low rates of decomposition. When plowed up and seeded to wheat, nitrates increased to a greater extent following alfalfa than following grasses for a period of 3 or 4 years, thus indicating the value of alfalfa as a source of available nitrogen.

The work of Newton et al. (1945) pointed out that rotations which included legumes and grasses, and the addition of barnyard manure, as





carried out for 25 to 30 years at Lethbridge, Lacombe and Indian Head Dominion Experimental Stations, reduced or prevented losses of organic matter and nitrogen, as compared to large losses which resulted from the equally old grain and fallow rotations at the same stations. Similar results have been obtained by Peevy et al. (op. cit.) in Iowa and on Gray Wooded soils in Alberta as presented by Toogood et al. (op. cit.).

#### Summarization of Soil Test Data

Parker et al. (op. cit.) have carried out summaries on both the Tennessee and North Carolina soil tests. Here, they established a low, medium, and high index for phosphorus, potassium, other nutrients, and pH on a county basis. Each nutrient was then mapped by shading in the counties according to the value of their nutrient or pH index. The general relationships of nutrient status to cropping and fertilizer practices and soil associations were discussed. The authors recognized that there may be considerable bias in connection with the data on soil samples taken and sent in by farmers. In addition, it was realized that the data summarized on a county basis are usually crossed and recrossed by soil areas and by different types of land use.

Tisdale and Nelson (op. cit.) show the work of Dumenil et al. in Iowa where levels of phosphorus and potassium have been shown in a map-like form of the various soil areas. A summary of the soil test data over crops grown illustrating the cropping effect on the level of nutrients in North Carolina was the work of Welch and Kamprath as portrayed by Tisdale and Nelson. Shickluna, as presented by Tisdale and Nelson, compared the soil texture to levels of potassium found in the soils tested as part of a





soil test summary. There are many types of summaries and techniques that can be used to summarize soil test data. There is, however, very little published literature on such summaries to illustrate the advantages and disadvantages of various methods.

A study to determine the accuracy of a soil test summary was conducted in North Carolina by McCollum and Nelson (1954). The results of their study are quite interesting. Their sampling was done systematically and the results compared to those of the voluntary farmer samples which had been received. The farmer samples tended to be a little higher in organic matter and pH, but not significantly so. The variation in pH might possibly be due to seasonal variation since the farmer samples were received in spring and the systematic sampling was carried out in the summer months. The high discrepancy between the calcium and magnesium results was attributed to fertilization practices. The farmers' samples tended to be higher in available potassium for all crops, significantly so for 7 of the 10 comparisons. This could not be tied to different rates of fertilization and thus it was believed that the depletion of the growing crop combined with the slight build-up of exchangeable potassium from fall to spring as reported by Rouse and Bertramson (1949) was most likely the cause. The data for phosphorus shows an almost complete reversal of the trend established for the other analyses; the systematic samples tended to be higher, but significance was indicated in only 4 comparisons.

It was the view of McCollum and Nelson that the errors associated with farmer sampling may be eliminated as a major consideration in summarizing soil test results, not because the farmers take valid samples,



but with large numbers of samples, that are well distributed over an area, there will be a reduction in the error due to random variation. It is my opinion that their proof for eliminating errors associated with farmer sampling is quite inadequate and a closer look is needed.

### Mapping Techniques

There are many techniques available for mapping and these techniques will vary considerably depending upon the nature of the data to be mapped, the importance of it being mapped, the accuracy required, the equipment and time available, and the preferences of the users. The following discussion will deal with some analytical techniques available that can be applied to a computer and mass amounts of numerical data collected from known locations over a specified area.

Regional and local trends: The problem of distinguishing between regional and local effects can be approached from an analytical point of view. Krumbein (1956) states that in this approach facies maps of any scale are considered as having both regional and local components, and the problem resolves itself to separation of the components by graphic or other methods. The analytical method is independent of map scale, size of stratigraphic unit, or density of control, although interpretation of results is influenced by these factors. Basic to this approach is the use of data expressed as numbers. The numerical data can be expressed on contour-type maps, which are two-dimensional representations of surfaces. The surfaces represented by the maps may be thought of as response surfaces generated



by geological factors and climatic factors. Some of these factors operate on a relatively large scale, and others on a relatively local scale. The larger-scale controls give rise to regional components in the response surface. The smaller-scale controls give rise to local components (sometimes called "residuals"). The observed response surface represented by a contour-type map of any scale thus reflects the simultaneous effect of widespread and local controls commensurate with scale of study.

The regional component in a facies map of any scale may be defined as that part of an observed value which is relatively stable and changes systematically if at all from point to point on the map. The regional component gives rise to a relatively smooth surface. In contrast, the local component is a relatively unstable part of the observed value that varies in a more or less irregular manner over the area. It contains no significant gradients and represents a series of non-systematic (positive or negative) departures from the smooth regional form. Thus, the local component by itself would display a spotty pattern of hills and hollows if it were mapped separately from the regional component.

When the control in a broad facies map is fairly open, the size of one grid cell may become larger than the entire map of a detailed area. Hence the terms "regional" and "local" take on a different scale of meaning. If a map is analyzed on a township basis then the more prominent local effects will be approximately township size. The regional surface will be the smooth surface that extends across several counties or even the whole mapped region.

Grant (1957) sees the word "regional" as one of those unfortunate





terms in the jargon of geophysics that has become irremediably equivocal by usage. In his opinion the word "trend" in this general context is more satisfactory, since it has a more precise meaning and - most importantly - is capable of special applications and a fairly rigorous definition.

"Eyeball" method: Krumbein (1956) states that one of the simplest methods for estimating the regional effect is to draw a series of smooth contours over the observed map to remove the irregularities and reversals in the facies lines. The smooth contours represent the regional facies surface, and the numerical difference between the observed and smoothed contours at any point represents the local facies effects. This rapid "eyeball" method is based on the principle that the response surface for the large-scale effects must by definition be a relatively smooth surface. Experience is a requirement for using the rapid method, inasmuch as considerable flexibility is possible in the smoothing process.

Profile method: A more formal method for separating the regional effect is based on a series of equally spaced profiles along and across the map. Each profile is reduced to a smooth curve, and by fitting the profiles together at each profile intersection, a set of internally consistent values is obtained on a square grid. Several adjustments are usually needed to obtain smooth profiles that are internally consistent and reasonable in both directions. Krumbein (1956) states that the grid need not be set parallel with land subdivisions. There are some advantages to aligning the grid with the facies trends when an area is selected for analysis from a larger map.





The degree of smoothing depends on subjective judgement, but in the smoothing process an attempt is made to keep the number of inflection points in the smooth lines to a minimum consistent with the data. In this way an undue number of secondary humps and hollows that might be local are not carried over to the regional surface.

Circle method: An alternative to the profile method involves adjusting the observed values at a series of grid points by bringing in the influence of near-by surrounding points. Griffin (1949) introduced the circle method where a circle of unit radius is drawn about each grid point on the original map and interpolated values are read at equal spacing on the circumference of the circle. The average of these points is then used as an adjusted central value. By repeating this process at all grid points, a series of regional values is obtained which may be contoured on the map.

Bayrock and Pawluk (1966) used a somewhat different averaging technique which they called a moving weighted averaging technique. The technique involved weighting more heavily those points closest to the central value, thus obtaining a weighted average for the adjusted central value. Using this technique, with the aid of the IBM 1640 and 1620 computers, they produced concentration maps for some trace elements found in Alberta tills.

"Expected value" method: Krumbein (1956) states that the "expected value" method is a much simplified method that should be used on maps that are relatively free from complexities that may be introduced by significant cross-product polynomials. In this method the observed grid values are arranged in rows and columns and a mean obtained for each. The "expected



value" at each grid point is obtained by the row mean plus the column mean minus the overall mean.

The "expected value" method is easily applied and requires only a slide rule and an adding machine. It has very severe shortcomings in complex situations, inasmuch as it takes no account of interaction terms that are sometimes significant when the underlying facies trends lie at an angle to the arbitrary grid direction. In addition, it automatically includes all direct polynomials in the regional surface, and hence it is not overly useful for large maps with many rows and columns. Despite its shortcomings, it does provide a quick method for rough analysis of limited areas within larger facies maps.

Abbreviated regression using orthogonal polynomials: The abbreviated regression using orthogonal polynomials is a shortened method of computation that can be used to evaluate the direct u and v polynomials from row and column sums. Krumbein (1956) was the first to introduce this method and the basic theory underlying the method. In this case a single observation is composed of a population mean, plus a contribution from the row, plus a contribution from the column, plus a residual contribution commonly called "error". By the use of a table of orthogonal polynomials as supplied by Anderson and Houseman (1942) or DeLury (1950) all or some of the direct u and v polynomials can be computed with the row and column totals. The abbreviated method neglects all cross-product terms in the regional surface but permits selection of the more important direct u and v polynomials. The "expected value" method retains in the regional surface all the direct u and v polynomials and assigns all cross-products



to the residual map. In general, the abbreviated method is suitable for fairly rapid analysis of maps with numerous rows and columns, in which the expected value method may add too many minor undulations to the regional surface. Computations for the abbreviated regression method are shown in Krumbein (1956).

Complete regression using orthogonal polynomials: Oldham and Sutherland (1955) published what may be considered the main paper on the use of orthogonal polynomials for mapping. Computational and theory details for the complete regression method are also given in Bennett and Franklin (1954), Anderson and Bancroft (1952), Davies (1954), Box (1954) and DeLury (op. cit.). Essentially the method consists of computing a series of polynomial coefficients and sums of squares for all direct polynomials and cross-product polynomials in the design. The sums of squares can be used to select the more important terms to be included in the regional surface. The complete regression method is used when the grid orientations or facies trends are such that it is not considered safe to neglect the cross-product polynomials in the analysis; moreover, the general pattern is such that it is desirable to segregate the more important direct and cross-product polynomials for construction of the regional surface.

The method of orthogonal polynomials requires that the experimental observations be known at equally spaced intervals. However, Krumbein (1959) and Whitten (1959) have developed an IBM 650 program for non-orthogonal polynomial analysis for trend surface analysis of contour-type maps with irregular control-point spacing. Both authors feel the program is





advantageous to geological data in "feeling out" the data in early stages of many map studies, to see whether some underlying tendencies may be of help in further study or exploration of an area.

Statistical analysis of facies maps have been discussed by Krumbein and Miller (1953), Krumbein (1955a), and Krumbein (1955b). The separation of the components into two sets, one representing the regional trend and the other the local effects, can usually be done without the aid of statistics. DeLury (op. cit.) suggests a test proposed by Bartlett may prove helpful in cases where the decision is difficult to reach.

Multiple regression method: The multiple regression method is very similar to the complete regression analysis using orthogonal polynomials. Walker, Hall and Protz (1968) have used this technique in summarizing soil trends and variability across three selected landscapes in Iowa. The method has a distinct advantage in that computer programs are readily available for this type of analysis.

#### Simulation and Prediction of Unknowns

Often, in soil testing, certain chemical measurements are read up to, but not beyond, a pre-decided limit. For purposes of recommendations this is fine, but for purposes of averaging and statistical summarization this is inadequate. These "ceiling values" will bias the results of a summary, especially if many of them are involved. There exists, then, a real need to obtain or otherwise predict the actual value or a reasonable approximation of that value. This can be achieved by simulation.





Webster's Dictionary defines simulation as the process of reproducing under test conditions phenomena likely to occur in actual performance. Flagle (1960), in defining simulation, explains that simple games fascinate people because they "simulate" some real experience. It is also possible, he continues, that this similarity exists in many other less dramatic but equally important matters. Churchman et al. (1957) defines a scientific model (i.e. that model used to simulate an event) as a representation of some subject of inquiry and used for purposes of prediction and control. Once the model is constructed, one can simulate the event.

Krumbein (1955b), in mapping sand-shale ratio values, found that these values followed a Gamma distribution. Through the use of random numbers he drew values from the model distribution and plotted them on a map. The final result showed a striking resemblance to that of the original map. A similar method could easily be applied to "ceiling values", where their probability distribution is known.

As a technique, simulation is remarkably useful and easily applied in a computer program. However, simulation of any event is no more accurate than the model used.

#### Summary of Literature Review

Over the years a large amount of soil test data has been collected. Shickluna and Walsh (1967) state that the efficient processing of soil test data is a prerequisite for the successful operation of any soil testing program. One of the more recent tools used to accomplish this is the electronic computer.



The soil testing program starts with the collection of a soil sample from a field. Generally, all the researchers agree on two points: the sampling area should be a homogeneous soil population with respect to soil type, plant growth, and treatment; and, accuracy can be increased as the number of cores composing the sample increases. The literature dealing with the number of cores required is quite variable. Some authors say from 15-20 on a ten acre field. Others suggest from 2 to 500 on less than an acre depending upon the precision required. According to one researcher, the degree of precision or accuracy of sampling is generally not affected by field size. Various statistical techniques to determine the number of cores required or the optimum type of sampling scheme to use are discussed in the available literature.

Several workers point out that phosphorus is the most difficult nutrient to sample because of its variability. It is therefore usually the least accurately determined measurement with respect to the number of cores used. Soil reaction, on the other hand, seems to be one of the least variable measurements. The number of cores required for any given degree of precision increases where chemical fertilizer or manure have been applied.

The level of a nutrient in the soil is influenced by cropping practice, fertilization, and sometimes season. It is generally assumed that nitrate nitrogen is most affected and therefore most variable with respect to the above three factors. Phosphorus tends to be low in grassland soils and high in market-garden soils where high applications of fertilizer are usually applied. Little is known of the effect of crop type on levels of potassium in the soil. Soil reaction is perhaps a tenth of a unit or so lower on continually cropped-land or where nitrogenous fertilizers have been applied.



Summaries of soil test data are now being carried out in many countries. However, very few of these summaries are published. Most summaries compare the nutrient levels to cropping practices and occasionally to soil management groups. No literature exists on the mapping of soil test data.

Many types of mapping techniques exist and the majority of these can be efficiently applied to the computer, especially where large amounts of data exist in numerical form. Two main categories of mapping methods exist: the averaging and the statistical. The averaging techniques are easy to apply to a large grid where areas exist with no samples, but cannot be used to describe the surface obtained. The statistical techniques are not so easy to apply and usually require that every grid-intersection have a value. However, they are advantageous in that the regional trend can rapidly be separated from the local effects and properly defined.

Computers are commonly used today to simulate an event and predict an outcome. Many researchers in economics and other fields frequently use models to reproduce a phenomena likely to occur in actual performance. Similarly, in soil testing, the computer can be used to simulate quickly the sampling of a field or to predict an unknown value by simulating its known distribution.

In conclusion, evidence available in the literature bears out the importance of the many factors that must be taken into account in the summarization and mapping of soil test data. In addition, it emphasizes the importance of the electronic computer as a significant tool in summarizing such data.



### III. MATERIALS AND METHODS

#### Materials

Soil test data: Since the establishment of the Alberta Soil and Feed Testing Laboratory in 1956, about 75,000 farm soil samples have been analyzed. The information available from these samples has been carefully filed and stored by the Laboratory as it was received. The amount of information accumulated has increased rapidly in the last few years as the numbers in Table I below clearly indicate.

Table I. Distribution of soil samples received by the A.S.F.T.L. based on date received.

<u>Year</u>	<u>No. of Samples Received</u>	<u>Percent of Total</u>
1962	2610	4
1963	4980	8
1964	5480	9
1965	6470	11
1966	12910	21
1967	15390	25
1968	13140	22

The farm soil samples make up the majority of samples analyzed by the Laboratory. However, considerable numbers of research, demonstration and greenhouse samples are also analyzed. The above tabulation and all results illustrated or discussed in this study included only the farm soil samples.







The data available for each sample consisted of farmer information and results of laboratory analyses. Details of information collected may be found in Appendix I. The key information supplied by the farmer and needed for summarization purposes included the following: the date of sampling, the number of cores representing the sample, the previous crop grown, the legal location and the type and number of district.

The Laboratory data included analyses for nitrate nitrogen, available phosphorus and potassium, soil reaction, and soluble salts, as well as qualitative analyses for sodium, sulphate, organic matter, lime, and texture. Those that are considered in this particular project included only the first four named. The procedures used in the analyses are presented briefly below.

(a) Preliminary - All samples were placed in a thermostatically controlled exhaust oven and dried at 65°C. The sample was then crushed to pass through a 2.0 mm sieve. This was accomplished by placing the sample in a Hewitt Soil Grinder with a 2.0 mm sieve.

(b) Nitrate nitrogen - The procedure was a photometric determination based on the work of Harper (1924) and Prince (1945). A solution of 0.02 N  $\text{CuSO}_4$  and 0.007 N  $\text{Ag}_2\text{SO}_4$  acted as the nitrate extracting solution. Phenoldisulphonic acid in the presence of  $\text{NH}_4\text{OH}$  produced a yellow colour due to the reaction with the nitrate nitrogen in the soil extract. The colour intensity was read on a Bausch and Lomb "Spectronic 20" colorimeter at wavelength 415 m $\mu$  using a flow through cuvette.

(c) Available phosphorus - The extracting solution consisted of 0.03 N  $\text{NH}_4\text{F}$  in 0.03 N sulfuric acid, which removed a proportion of the more available soil phosphorus. A combined nitric acid-vanadate-molybdate solution was used for colour development. A "Spectronic 20" at wavelength



400 mu was used to measure the colour intensity.

(d) Available potassium - The potassium extracting solution was made up of 1N ammonium acetate. Lithium nitrate solution was added to the filtrate and the potassium content determined by the flame photometer.

(e) Soil reaction - The pH values were determined in a 1:1 soil paste using a pH meter equipped with a glass and calomel electrode. The procedure is outlined by Doughty (1941).

Except for soil reaction, all the above analyses were recorded in pounds per acre.

The Laboratory also assigned a soil-climatic zone to the sample based on its legal location. These soil-climatic zones are based mainly on the work of Odynsky (1962) and Bowser (1967). The zones or soil areas were used for making recommendations and for summarization of the data. Figure 1a, on the following page is a map of Alberta showing the soil-climatic areas as designated by the Soil Testing Laboratory. Figure 1b is a map of Alberta showing the Counties, Municipal Districts, Improvement Districts, and Special Areas.

These maps will be referred to later when some summarizations by soil area and county are shown. The relationship of the soil areas the Laboratory assigned to the Soil Zones of Alberta is as follows:

- Area 1 - - Brown soil zone.
- Area 2 - - Dark Brown soil zone south of the Bow River.
- Area 3 - - Dark Brown soil zone north of the Bow River.
- Area 4 - - Thin Black soil zone south of the Bow River.
- Area 5 - - Thin Black soil zone north of the Bow River.
- Area 6 - - Black and Dark Gray soil zones west of range 20, W4.



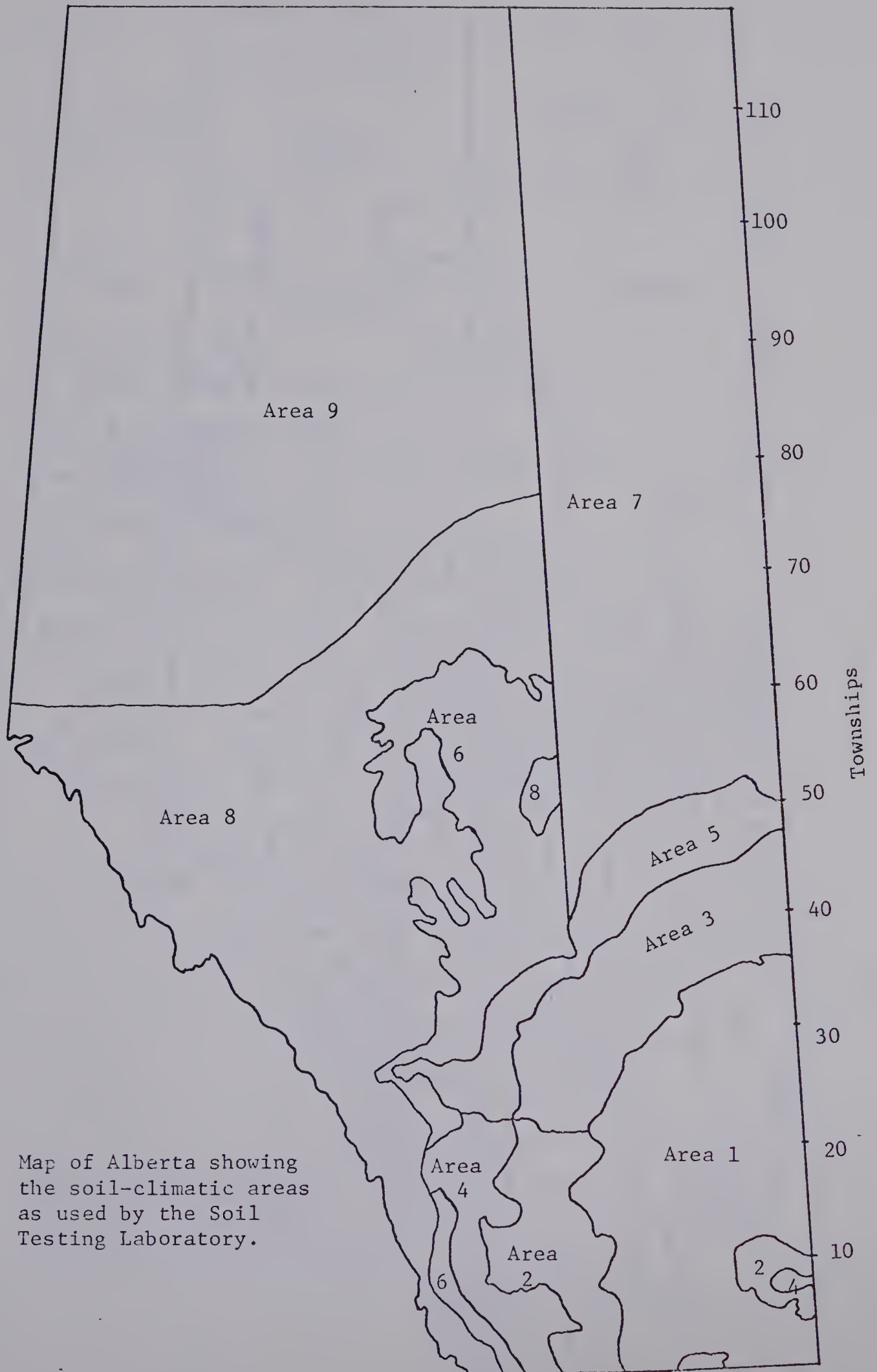
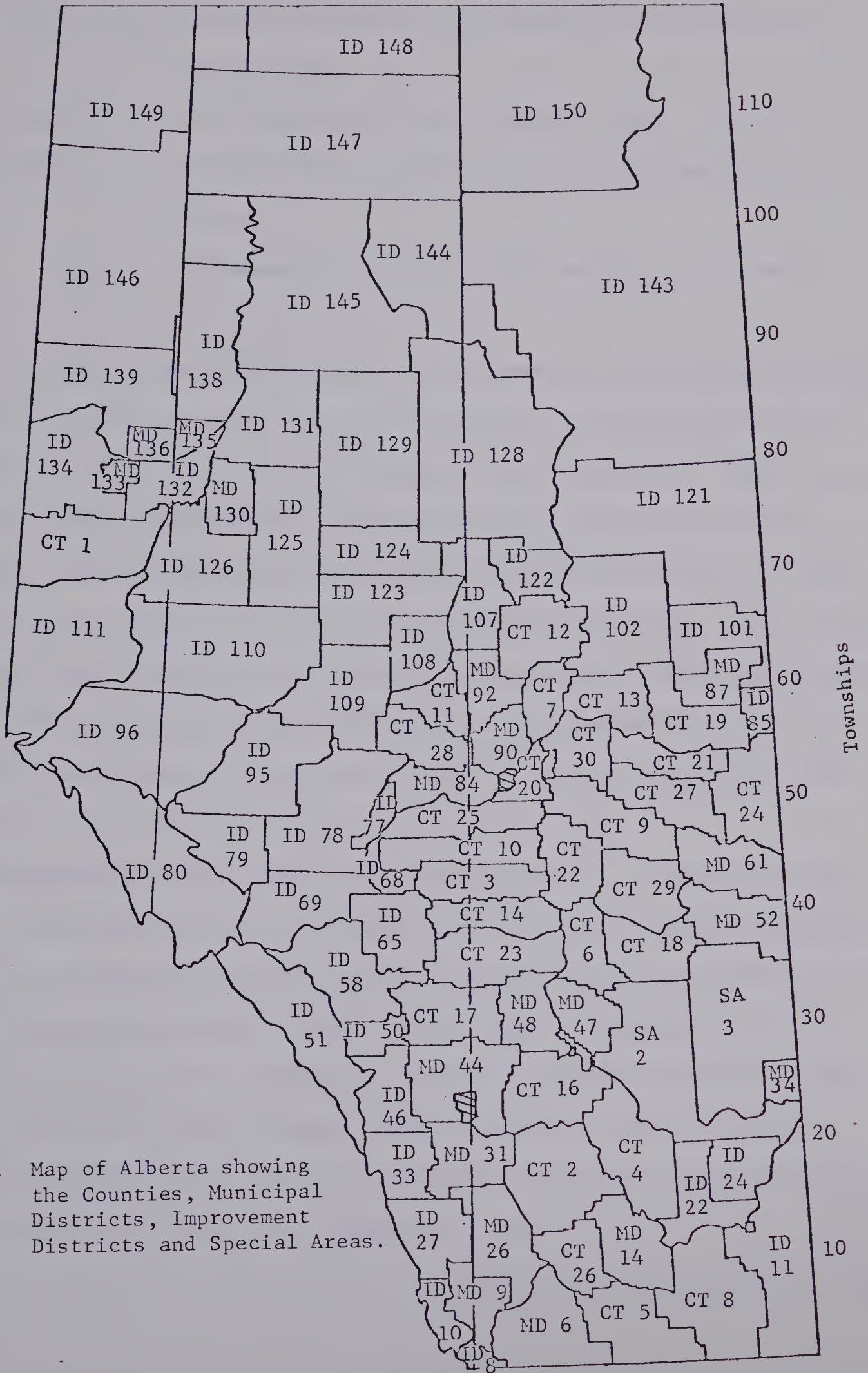


Figure 1 a. Map of Alberta showing the soil-climatic areas as used by the Soil Testing Laboratory.









Area 7 - - Black, Dark Gray and Gray Wooded soil zones east of range 21, W4.

Area 8 - - Gray Wooded soils west of Areas 6 and 7.

Area 9 - - Dark Gray and Gray Wooded soils of the Peace River region.

Area 10 - - Irrigated soils located mainly within Areas 1 and 2.

With the accumulation of such a large amount of data there were bound to be a few discrepancies. The data supplied by the farmer were viewed critically. The accuracy of the data was open to question. Often farmers tended to give an estimated or rounded-off value for yield, fertilizer applied, number of cores representing the sample, and field size. Sometimes farmers did not complete all the required information. Legal location did not always check against district type and number. In addition, during the past few years Improvement Districts have become Municipal Districts and Municipal Districts have become Counties. These changes have to be noted.

Discrepancies also exist in the Laboratory data. Nitrate nitrogen, until quite recently was only measured up to 80 lb/ac. The new potassium test was introduced on April 1, 1965 (Laboratory Report No. B-4728). However the results were only read up to 600 lb/ac. On November 11, 1967 (Laboratory Report No. B-12806) the Laboratory began recording the results up to 800 lb/ac. Then in August, 1968 (Laboratory Report No. B-15750) a maximum reading of 998 lb/ac was established. Phosphorus was recorded up to 257 lb/ac and very rarely exceeded this value.

In addition, it was noted that the results as recorded were not really



continuous. This was especially true at the higher values where the colorimeter or flame photometer could not be read as accurately as at the lower values where a wider scale existed. Such a situation resulted in the recurrence of certain values. However, statistically such data were considered as continuous.

Cameron farm data: The Cameron farm data were collected and analyzed mainly for the purpose of studying the accuracy obtained (when randomly sampling a field) with a varying number of cores making up the composite. The field selected was a summerfallow field located in the northern portion of SE 25-56-27-W4 (see Figure 2). This is about 30 miles north-west of Edmonton. The field was 30 acres in size, rectangular in shape (2120 x 600 feet), with the length running east-west. The sampling of the field was carried out using a tube-type sampler on the 5th and 6th of September, 1968. Individual cores were taken approximately every 40 feet across the field and every 46 feet along the length of the field. A total of 690 cores were taken and analyzed individually. Each core was 1 inch in diameter and 6 inches deep.

In 1966 the field was seeded to barley with 16-20-0 applied at 85 lb/ac. In 1967 a crop of mixed oats and barley was grown with 16-20-0 applied at 60 lb/ac. In 1968 the field was summerfallowed. According to the Soil Survey of the Edmonton Sheet (Bowser et al., 1962) the soil is classified as 50% Falun Loam (Orthic Dark Gray), 40% Angus Ridge Loam (Eluviated Black), and 10% Kavanagh Loam (Black and Dark Gray Solodized Solonetz and Solonetz). Approximately 90% is developed on glacial till and the remainder on residual material. The topography is undulating in the east half to gently rolling in the west half.



Trial mapping data: An area was chosen in Alberta for trial runs on the various mapping techniques. The area extended from range 12 to 25, W4 and township 32 to 57 as outlined in Figure 2 on the following page. A border of 5 townships about the mapped area was allowed so the edge effect would be eliminated. A border of only two townships existed on the west side next to the 5th meridian. Only those soil samples submitted in the fall 1967 were used for this particular study.

The area was selected because it had in the past been an area where there was a sufficient number of samples sent in to be analyzed. The element chosen for mapping was phosphorus.

All the sample data from the area chosen and the accompanying border townships were transferred to a separate magnetic tape in a shortened format for more rapid use. Samples were eliminated where an area greater than 10 acres was represented by one core and where organic soils existed. The data were used to test out the mapping techniques.

Other data: Other data included small amounts of data that were analyzed separately to aid in the completion of the project. In the fall of 1966, the A.S.F.T.L. carried out a Nitrate Survey. Some of the results of this project were used for comparing the nutrient levels of stubble and adjacent fallow fields. Although deeper samples were taken, only the results for the top 6 inches of soil were used. A total of 45 different farms was used, each having adjacent stubble and fallow fields. These farms were widely spread over Alberta. About 10 acres of each field were sampled with a minimum of 10 cores making up the composite. In the spring of 1967, thirteen of the original 45 farms were resampled. This added











sampling allowed a check on spring versus fall sampling.

In the spring of 1968 the A.S.F.T.L. carried out an analysis for potassium on 140 soil samples in which the potassium content was known to be greater than 600 lb/ac. Later, 73 of these samples were used to derive soil area distributions for potassium content in excess of 1000 lb/ac. In the fall of 1968 the Laboratory analyzed 387 soil samples for nitrate nitrogen. The samples analyzed had previously been recorded as 80 + lb/ac. These results were used to derive a distribution for nitrate nitrogen in excess of 80 lb/ac.

A brief check was required to test whether there was a variation in soil nutrient status from year to year where no change occurred in fertilizer and cropping practice. Part of the data collected at the six Economic-Productivity sites in the Edmonton and Lacombe area was used for this study. The treatments selected and years compared are shown below.

<u>Barley Treatments</u>	<u>N</u>	<u>P</u>	<u>K</u>
7	90	36	12
8	90	36	36
21	120	48	0
22	120	48	48

<u>Forage Treatments</u>	<u>N</u>	<u>P</u>	<u>K</u>
7	120	45	12
8	120	45	36
9	80	30	24
15	80	30	48



<u>Site</u>	<u>Barley Years Compared</u>	<u>Forage Years Compared</u>
1 and 2	65,66,67	66,67
3 and 4	66,67	
5 and 6	66,67	66,67
21 and 22	65,66,67	66,67
23 and 24	66,67	
25 and 26	66,67	66,67

The data used were the results of analyses for nitrogen, phosphorus, potassium, and pH only in the top six inches. The samples analyzed were subsamples from composites of 4 thoroughly mixed cores per replicate. These samples were all taken in spring before seeding and fertilizing.

### Methods

Methods are divided into two sections: A. data processing, and B. programming. The programming phase is further subdivided into various categories in which the techniques used are emphasized in the discussion rather than the actual details of programming.

#### A. Data Processing

Data processing in this project refers to the transferal of data from written or typed form to a form readily available for computer use. The exact point where data processing stops and programming starts is not always distinctly defined. However, in the case of this project, the processing was considered to be completed once the magnetic tape containing the information had been created.



A preliminary to the data processing was the development of a format suitable to handle the data needed. The format developed for the A.S.F.T.L. records consisted of 324 characters per sample (record) and required five IBM punch cards. The format is illustrated in Appendix I.

The data were transferred to IBM computer cards by trained operators using IBM 024 and 026 Card Punches, and an IBM 056 Verifier to check the punched cards. The data on the cards were transferred to magnetic tape and sorted and merged. A copy of this tape was then obtained for use on the IBM 360/67 computer located in the Department of Computing Science at the University of Alberta.

As of March, 1969 there were 62,570 sample records (dating back to April, 1962) available on tape.

In addition to the transferal of soil test reports to cards and then to tape, other research data were punched on cards. These included the Cameron farm data, fallow vs. stubble, spring vs. fall, and the high potassium and nitrate nitrogen distribution data.

## B. Programming

All of the results of this thesis have been obtained by means of computer programs written in Fortran IV. To discuss thoroughly the programs used would be time consuming and unnecessary, for there were over 70 programs written to analyze the data. The following is a discussion of the major techniques or methods that were incorporated into the programs.

Sampling techniques: The data collected from the Cameron farm were used to check the reliability of sampling a field. Two approaches have been used to analyze these data. The first approach was statistical and involved the



use of Stein's two-stage sampling formula as presented by Guenther (1965). The formula is shown below:

$$n = \left| \frac{2 t_{1-\alpha/2}}{L} \right|^2 x s^2$$

where,  $n$  = number of cores that should be taken

$t$  = value from  $t$ -distribution table depending upon the degrees of freedom and alpha ( $\alpha$ ) level.

$\alpha$  = level of significance - determines the confidence interval one wishes to establish.

$L$  = total length of the interval about the mean, i.e., the level of precision one wishes to obtain. This may be expressed as a percentage of the mean.

$s^2$  = variance of the preliminary sample taken.

The procedure involved calculating the field variance for the four determinations studied. The calculated variance for each determination was applied to the formula with varying confidence limits and levels of precision. The confidence limits used ranged from 5% to 99%. The level of precision, as determined by  $L$  in the formula, varied depending upon the particular chemical analysis under study. The level of precision ranged from 2 to 20 lb/ac for nitrate nitrogen, 2 to 40 lb/ac for phosphorus, 10 to 170 lb/ac for potassium, and 0.2 to 2.0 units for pH. A measure of the percent accuracy was obtained directly from the confidence limits at each of the levels of precision for 2, 4, 6, . . . , and 40 cores.

The second approach involved a simulation of actual field sampling.





Random sampling without replacement, using a random number generator, was simulated 10,000 times for 2, 4, 6, ..., and 40 cores. The calculated mean for each combination of cores selected was compared to that of the field mean and grouped according to its precision level or closeness to the field mean. A percent accuracy was then determined for each precision grouping with respect to the number of cores used. The results obtained were compared to those of the statistical approach.

Stein's formula was also used to calculate the number of cores required for different acre-sizes within the field. The acre-sizes decided upon were 0.5, 1, 5, 10, 15, 20, and 25. The procedure involved the movement of a particular acre-size throughout the field, never in the same location more than once. For each movement, the variance was calculated, followed by the calculation of the number of cores required. The final result involved 693, 684, 445, 142, 152, 114, and 60 individual movements for each acre-size respectively. The average number of cores required and the range were recorded for the 70%, 80%, 90%, and 95% confidence-limits and the desired levels of precision. An examination was made on the results to determine the effect of field size on the number of cores required.

The mapping technique selected for the province has, within it, a critical circle which is used to discontinue the calculation of a map value if the number of samples within this circle is too low for a reliable calculation. Bayrock and Pawluk (1967) refer to this number as a weighting number. A more complete discussion on the critical circle may be found in Appendix III and a diagram may be found in Appendix IIb. It was felt necessary for this project to obtain an indication of the number



of samples required within a critical circle to make a reliable calculation for a final map value. This was achieved by the use of the sampling formula for a critical circle of 37 townships in size. Calculations were also carried out by the same procedure for the number of samples required within individual townships. In both cases an 80% confidence limit was used with varying levels of precision.

The procedure involved a revision of the Alberta mapping program so that the variance could be calculated within the critical circle or within individual townships. This particular technique allowed considerable control over certain variables that influence the variation. These included the time period covered, the crops from which the samples were selected, and the elimination of inadequate samples. In addition, the method required that 6 samples be present in a critical circle and 4 in a township before the sampling formula was applied.

The method described above was applied to the soil test data (by means of a computer program) and results were obtained for the plausible variations. The results recorded included the average number of samples required, the standard deviation, the percentage distribution of the number required, and the number of critical circles or townships that were used to derive the above.

The actual averages and distributions of the number of samples available in the critical circle and the individual townships for each of the above variations were also derived using a similar program.



Summarization techniques: A number of methods of summarizing soil test data have been presented by Tisdale and Nelson (op. cit.) and Parker et al. (op. cit.). A similar technique to that described by Parker et al. was used to summarize the A.S.F.T.L. data by County, Municipal District, Improvement District, and Special Area; but arithmetic means, standard deviations, and number of samples were calculated rather than an index value. This summarization technique working on the basis of districts was further developed to separate the results according to four major crop categories: summerfallow, stubble, grass-legume, and breaking and vegetables. The incorporation of the technique into program form allowed control of the time period over which the summary could be conducted and made the elimination of inadequate samples an easy task.

A very similar method was employed in summarizing the data by the ten Laboratory designated soil-climatic areas.

The programs utilizing these techniques were developed to carry out summaries for soil nitrogen, phosphorus, potassium and pH simultaneously. All calculations on pH were computed using pH values, rather than converting back to hydrogen ion concentration. Shine and Chin (1957) found that pH values follow a normal distribution, whereas values of hydrogen ion concentration rarely do.

Other less important summarizations were completed to display distributions and arithmetic averages for field size, number of cores taken, year-month sample numbers and area-crop sample numbers.

Mapping techniques: The trial mapping runs were made on an area of Alberta previously described in the section "Materials" and as previously outlined in Figure 2. The purpose was to determine the most accurate and applicable





technique for the soil test data.

Before entering into the discussion it should be noted that the word "prediction" will be used occasionally. It is used to mean the arithmetic regional map value as calculated by the particular mapping function applied. In a sense the calculation actually does perform a scientific prediction, in that the derived map value is previously unknown.

The area in which the mapping techniques were tried was automatically arranged in rows and columns by the legal location values of the samples. The township numbers corresponded to row numbers and the range numbers corresponded to column numbers. The result was a "pseudo-grid" because of the improper alignment of the townships along the correction lines. However, for this study it was assumed to be a real grid with each township representing a grid-point.

(a) "Eyeballing" - - "Eyeballing" was not really a computer applied technique, however, a computer was used to speed up the presentation of the data for "eyeballing". The method involved simple averaging and rounding off the observed values for the individual townships. The results were listed in a grid-like representation of the area.

(b) Profile - The profile method, as described by Krumbein (1956), was not used but rather a technique which had a similar outcome and was easily adapted to a computer program. This was the use of a weighted moving average commonly used to smooth time-series curves and described by Hoel (1966).

The procedure involved moving a weighted averaging function along the rows and columns formed by the townships. The averaging function smoothed the map surface at each township by simultaneously averaging the observed





values of those townships directly to the north-south and east-west direction with that of the township in which the prediction was being made. The averaging technique permitted weights to be assigned to townships depending upon the distance from the central township and the degree of smoothness wanted. A diagram of the type of averaging function used and brief description of the computer application of the profile technique is given in Appendix IIa.

(c) Circle - - The circle method was based on the same averaging technique as the profile method. Instead of weighting and averaging for just the townships in the x- and y- direction however, the technique was amplified to include concentric rings of townships about the cell to be mapped. The technique involved taking into account all values within the main circle and weighting these values according to the weights assigned to the concentric circles encompassing them. This method allowed considerable control in the degree of smoothing of the map surface.

A diagram of the type of circle function used and a brief description of the computer application of this technique may be found in Appendix IIb. A fuller account of the control and application of the technique to the mapping of soil test values for the province may be found in Appendix III.

(d) "Expected value" - - The "expected value" method was applied to the 26 rows and 14 columns of the trial mapping area using the same procedure as described by Krumbein (1956). A prediction was made at each grid-point by the addition of the row and column mean and subsequent subtraction of the overall mean.

(e) Orthogonal polynomials - abbreviated method - - The abbreviated regression method using orthogonal polynomials, as described by Krumbein



(1956) and developed by Tukey (Krumbein, 1956), was applied to the trial mapping area using polynomial coefficients obtained from DeLury (1950) for a 14 x 14 and a 26 x 26 matrix. Townships having no samples were predicted using the circle method. Averages per township were used.

The method displayed the  $Z^2$  matrix which was used to select the direct polynomial coefficients accounting for the main effects. A revision of the technique permitted selection of standard coefficients and mapping of the results.

(f) Orthogonal polynomials - complete method -- The complete regression method using orthogonal polynomials, as presented by Oldham and Sutherland (1955), was applied to the map area using orthogonal polynomial coefficients obtained from DeLury (1950). The technique allowed the selection of any direct and interaction polynomial coefficients accounting for the main effects.

(g) Multiple regression analysis - - The technique involved fitting a polynomial to the mapping area by the use of a step-wise multiple regression program. A similar method to that described by Walker et al. (op. cit.) was used.

Alberta mapping technique: The procedure utilized to map the soil test levels for the province was based on the same basic principles as the procedure used by Bayrock and Pawluk (1967). Although, the function of the two methods was the same, the methods were substantially different, mainly because of the quantity of data and computing facilities available.

The procedure used makes use of the circle method with its critical circle, critical number, and controlled weighting effect. This mapping function is moved over a grid-like representation of the province formed by



legal location values. The mapped values obtained are plotted by a Calcomp Plotter only if enough samples are present within the critical circle to make a reliable prediction. The plotted map is a polyconic projection centered about the 5th meridian with coded numbers representing areas of high and low values.

A detailed write-up of the mapping program along with a copy of the program is available in the Department of Soil Science, University of Alberta. A shorter description of the Alberta mapping program and the circular mapping technique it employs may be found in Appendix III.

County and field mapping: The procedure followed for county and field mapping was very similar to that of the Alberta mapping technique. The circular weighted moving average was used as the mapping function.

The soil test data from townships 39 to 51 and ranges 15 to 23, W4 were used for the study of mapping on a county basis. This area included the whole of the County of Camrose. The mapping unit used, however, was a section rather than a township. Thus the only significant alteration of the mapping procedure was the conversion of legal location values into sectional-type grid cells and subsequent alignment of these cells into a grid-like pattern.

Mapping on the field basis entailed the use of the Cameron farm data with the individual core values acting as observed grid-point values.

Revision and prediction of unknowns: Revision and shortening of the original format was necessary for three main reasons: 1. the old format required unnecessary reading time, 2. a revision allowed the elimination of errors.





existing on the old tape, and 3. a revision allowed the simulation of previously unknown values for potassium and nitrogen. Of these three, only the last one required considerable preparation and methodology

Probability distributions for values of potassium over 600 lb/ac, 800 lb/ac, and 998 lb/ac for each soil-climatic area were prepared using special analyses carried out by the Laboratory. Similarly, a probability distribution was prepared for nitrate nitrogen over 80 lb/ac using the results of the 383 high nitrate analyses performed by the Laboratory.

The distributions derived formed model distributions which served as the basis for the computer program constructed. With the aid of a random number generator a value for every "ceiling value" could be simulated. In every case the simulated value could be classified as a "feasible" or "probable" or "most likely" occurrence for the actual distribution. The simulated value replaced the "ceiling value" in the revised format of the data.

Other techniques: Other techniques were incorporated into program form to analyze specific parts of the available data. These included an analysis - of-variance with Duncan's multiple range test and the paired and unpaired t-tests as described in Steel and Torrie (op. cit.). As the need arose, other techniques were developed to derive distributions, check results, and investigate new information.





#### IV. RESULTS AND DISCUSSION

Computer summarization and mapping of 62,500 soil test samples can produce large amounts of useful information. However, due to lack of space, not all the information obtained can be presented. Thus, the main discussion will deal with a few selected results pertaining to this study. No attempt will be made to present all the material obtained.

The A.S.F.T.L. data used for this project are available on two magnetic tapes: one located at the Soil Test Laboratory and the other in the Department of Soil Science, University of Alberta. Similarly all the major results such as area and district summaries and plotted maps are available at the above two locations. The Cameron farm data and the other data used are documented and filed in the Department of Soil Science.

##### Soil Sampling

Field basis: The A.S.F.T.L. Information Sheet for Farm Soils has a space for the farmer to write the number of places sampled and the size in acres of his field. Such information, if valid, should give an indication of the accuracy of the sample obtained. This information was analyzed for the whole province using all the samples from 1962 to 1969. A summary of the information obtained for the number of cores taken, the field size, and the acres per core is shown in Tables II, III, and IV, in that order.

An examination of the tables shows that the mean number of cores taken to compose the farm soil sample is 16. This is a little lower than the suggested number by most researchers in the literature reviewed. About



Table II. Cumulative percentage distribution of number of cores taken to compose a farm soil sample. A.S.F.T.L. data, 1962-68.

Average number of cores taken - - - - - 16  
 Standard deviation - - - - - 10  
 Number of samples with available data - - - 56392

<u>Number of cores</u>	<u>Percentage</u>	<u>Number of cores</u>	<u>Percentage</u>
* .GE. 1	100	.GE. 21	18
.GE. 2	99	.GE. 22	17
.GE. 3	98	.GE. 23	16
.GE. 4	95	.GE. 24	16
.GE. 5	92	.GE. 25	15
.GE. 6	89	.GE. 26	9
.GE. 7	84	.GE. 27	9
.GE. 8	83	.GE. 28	9
.GE. 9	78	.GE. 29	8
.GE. 10	78	.GE. 30	8
.GE. 11	66	.GE. 31	5
.GE. 12	65	.GE. 32	5
.GE. 13	59	.GE. 33	4
.GE. 14	58	.GE. 34	4
.GE. 15	57	.GE. 35	4
.GE. 16	42	.GE. 36	3
.GE. 17	40	.GE. 37	3
.GE. 18	39	.GE. 38	3
.GE. 19	36	.GE. 39	3
.GE. 20	36	.GE. 40	3

\* .GE. - - Greater than or equal to



Table III. Cumulative percentage distribution of field size.  
A.S.F.T.L. data, 1962-69.

Average field size - - - - - 57  
Standard deviation - - - - - 49  
Number of sample with available data - - - 56955

<u>Number of acres</u>	<u>Percentage</u>	<u>Number of acres</u>	<u>Percentage</u>
* .LE. 5	1	.LE. 55	64
.LE. 10	5	.LE. 60	70
.LE. 15	9	.LE. 65	72
.LE. 20	17	.LE. 70	75
.LE. 25	23	.LE. 75	77
.LE. 30	31	.LE. 80	84
.LE. 35	36	.LE. 85	85
.LE. 40	50	.LE. 90	86
.LE. 45	53	.LE. 95	86
.LE. 50	62	.LE. 100	90

\* .LE. - - Less than or equal to

Table IV. Cumulative percentage distribution of acres per core.  
A.S.F.T.L. data, 1962-69.

Average acres/core - - - - - 6  
Standard deviation - - - - - 13  
Number of samples with available data - - - 55264

<u>Acres per core</u>	<u>Percentage</u>	<u>Acres per core</u>	<u>Percentage</u>
* .LE. 1	7	.LE. 11	92
.LE. 2	31	.LE. 12	93
.LE. 3	51	.LE. 13	93
.LE. 4	65	.LE. 14	94
.LE. 5	75	.LE. 15	95
.LE. 6	79	.LE. 16	96
.LE. 7	83	.LE. 17	96
.LE. 8	87	.LE. 18	96
.LE. 9	88	.LE. 19	96
.LE. 10	91	.LE. 20	97

\* .LE. - - Less than or equal to



80% of the samples are composed of at least 9 cores. The field sizes (57 acres) are much larger on the average than the 10 acres usually recommended. In fact, 95% of the fields sampled are over 10 acres. The average number of acres represented by one core is 6 with 80% of the samples having less than 7 acres represented by one core. Approximately 12% of the farm reports sent in are missing data on either the number of places sampled or field size or both.

A study of just the data received in the fall of 1968 indicates a change in the sampling scheme. The results show an increase in the number of cores as well as the field size. On the average, 18 cores are used to compose a soil sample from an average field size of 62 acres. The average acres per core is 4 with 80% of the samples having less than 5 acres per core. This appears to be a very good improvement over the previous summary representing all of the data. In addition, the fall 1968 data indicate that 80% of the farmers sent in samples having at least 10 cores per sample and from a field of less than 80 acres. The trend for more cores to be taken is possibly due to the increased sampling by commercial companies with modern, rapid sampling methods and to a better education, in general, as to the importance of an adequate soil sample.

The reports in the literature on sampling often fail to indicate the accuracy or the ability of specified numbers of cores to represent a field. Often the chemical analysis used for a particular test (for a particular nutrient) is not the same as that used by other researchers or soil testing laboratories. The result is often confusion. Therefore it was thought that a study along these lines could be used to determine the actual effect of the number of cores on the accuracy of a sample and, furthermore, the results obtained could be related to the available





literature and the A.S.F.T.L. data. This need led to the intensive sampling of a representative farm field - - the acquisition of the Cameron farm data and subsequent analyses.

The Cameron farm data was analyzed both statistically and by simulation techniques to determine the effect of the number of cores on the accuracy of the resulting sample. The results by both methods for the percent accuracy of nitrogen, phosphorus, potassium, and pH sampling are shown in Tables V, VI, VII, and VIII, respectively. On comparing the results from the statistical analysis with the results obtained from the use of simulation techniques, one finds a very close similarity in all four Laboratory determinations. This is a very good indication that all four measurements follow a normal distribution, and likewise, that common statistical procedures can be used on such data without the uncertainty sometimes accompanying statistical analyses. The tables show that there is very little difference in the results of the two methods used, however, if the populations were not normal, the results by simulation would be the most valid.

The disadvantage of such results is that they are valid only for the field sampled and whether they can be inferred to represent a majority of the fields sampled in the province is unknown. However, it is believed by the author, that further work along these lines would most likely yield very similar results for other fields.

Tables V, VI, VII, and VIII show the relationship between the percent accuracy achieved with a given number of cores and with a given interval or level of precision. The A.S.F.T.L. data indicates that the long-term average is 16 cores composing a single sample. This gives a 91% accuracy



Table V. Percent accuracy of nitrogen sampling on the Cameron farm data as determined by simulation and statistical techniques.

Mean nitrate nitrogen value for field = 17.5 lb/ac.

		Interval ( $\pm$ lb/ac from the mean)						
		$\pm 1$	$\pm 2$	$\pm 3$	$\pm 4$	$\pm 5$	$\pm 6$	$\pm 7$
<u>Number of Cores</u>								
2	* Sm	9	19	28	37	46	54	61
	** St	8	15	25	31	40	51	56
4	Sm	13	26	38	49	60	69	77
	St	12	22	35	46	55	67	71
6	Sm	16	31	46	58	69	78	85
	St	15	30	43	55	65	76	80
8	Sm	18	35	50	64	75	84	90
	St	17	33	50	62	72	81	87
10	Sm	20	40	56	70	81	88	93
	St	19	38	54	67	77	85	91
12	Sm	22	41	60	75	85	91	95
	St	21	41	57	72	82	90	94
14	Sm	23	45	64	78	88	93	96
	St	23	44	62	75	85	92	96
16	Sm	25	48	70	81	91	95	98
	St	24	46	65	79	87	94	97
18	Sm	27	51	70	84	92	96	98
	St	26	50	69	81	90	96	98
20	Sm	28	54	73	86	93	97	99
	St	27	52	71	84	92	97	98
22	Sm	30	55	75	88	94	98	99
	St	28	54	73	86	93	97	99
24	Sm	30	57	77	89	95	98	99
	St	30	56	75	87	94	98	99

(Continued)



<u>Number of Cores</u>		<u>+ 1</u>	<u>+ 2</u>	<u>+ 3</u>	<u>+ 4</u>	<u>+ 5</u>	<u>+ 6</u>	<u>+ 7</u>
26	Sm	33	60	79	91	97	98	99
	St	31	57	77	89	95	98	100
28	Sm	32	60	80	92	97	99	100
	St	32	59	79	91	96	99	100
30	Sm	35	64	82	93	97	99	100
	St	33	61	80	92	97	99	100
32	Sm	35	64	84	94	98	99	100
	St	34	63	81	93	97	99	100
34	Sm	36	66	85	94	98	100	100
	St	35	65	83	93	98	100	100
36	Sm	38	68	86	95	99	100	100
	St	36	67	84	94	98	100	100
38	Sm	38	68	87	96	99	100	100
	St	37	68	86	94	98	100	100
40	Sm	39	70	88	96	99	100	100
	St	38	69	87	95	99	100	100

\* Simulation

\*\* Statistical

---

Table VI. Percent accuracy of phosphorus sampling on the Cameron farm data as determined by simulation and statistical techniques.

Mean phosphorus value for field = 50.3 lb/ac

		<u>Interval (<math>\pm</math> lb/ac from the mean)</u>						
<u>Number of Cores</u>		<u>+ 1</u>	<u>+ 3</u>	<u>+ 5</u>	<u>+ 7</u>	<u>+ 8</u>	<u>+10</u>	<u>+12</u>
2	* Sm	3	10	17	23	26	33	39
	** St	3	10	15	25	29	35	40
4	Sm	5	15	25	34	39	48	56
	St	5	15	25	35	41	45	55

(Continued)



<u>Number of Cores</u>		<u>± 1</u>	<u>± 3</u>	<u>± 5</u>	<u>± 7</u>	<u>± 8</u>	<u>±10</u>	<u>±12</u>
6	Sm	7	19	31	43	48	58	67
	St	6	19	30	40	45	55	65
8	Sm	8	22	36	48	54	65	74
	St	7	21	35	47	52	65	72
10	Sm	8	25	41	54	60	71	79
	St	8	24	39	52	57	70	77
12	Sm	9	27	43	57	64	75	83
	St	9	26	42	56	62	74	81
14	Sm	10	28	46	61	68	79	87
	St	10	28	45	60	66	77	85
16	Sm	10	31	49	65	71	82	89
	St	10	30	48	64	70	80	88
18	Sm	11	33	52	68	75	84	92
	St	11	32	51	67	72	83	90
20	Sm	12	34	54	70	76	86	93
	St	12	34	53	69	75	85	92
22	Sm	12	37	57	73	80	88	94
	St	12	36	55	71	77	87	93
24	Sm	13	39	60	76	82	90	95
	St	13	37	57	73	79	89	94
26	Sm	13	39	61	77	83	91	96
	St	13	38	59	75	81	90	95
28	Sm	14	41	62	79	85	93	97
	St	14	39	61	77	83	91	96
30	Sm	14	41	63	80	86	93	97
	St	14	40	63	79	85	92	97
32	Sm	14	43	65	81	88	94	98
	St	15	42	65	81	86	93	97
34	Sm	15	44	67	83	88	95	98
	St	15	44	67	82	88	94	98
36	Sm	16	46	69	84	90	96	98
	St	15	45	68	83	89	95	98

(Continued)





<u>Number of Cores</u>		<u>+ 1</u>	<u>+ 3</u>	<u>+ 5</u>	<u>+ 7</u>	<u>+ 8</u>	<u>+10</u>	<u>+12</u>
38	Sm	16	47	70	86	91	96	99
	St	16	45	69	84	89	95	98
40	Sm	17	48	72	87	91	97	99
	St	16	46	70	85	90	96	99
* Simulation								
** Statistical								

Table VII. Percent accuracy of potassium sampling on the Cameron farm data as determined by simulation and statistical techniques.

Mean potassium value for field = 331 lb/ac

Interval ( $\pm$  lb/ac from the mean)

<u>Number of Cores</u>		<u>+10</u>	<u>+20</u>	<u>+25</u>	<u>+35</u>	<u>+40</u>	<u>+50</u>	<u>+60</u>
2	* Sm	10	20	24	33	38	46	54
	** St	8	20	26	31	35	45	55
4	Sm	14	27	34	46	52	63	72
	St	12	25	33	45	50	60	70
6	Sm	18	34	42	56	62	73	82
	St	16	32	40	55	60	70	80
8	Sm	20	39	48	63	70	81	88
	St	20	37	47	60	67	77	87
10	Sm	23	43	52	68	75	86	92
	St	22	42	51	67	72	82	90
12	Sm	25	47	57	73	80	89	94
	St	24	46	55	71	77	86	94
14	Sm	27	52	62	78	84	92	96
	St	26	50	60	75	81	90	96
16	Sm	29	54	64	80	86	93	97
	St	28	53	64	79	85	92	97
18	Sm	30	57	67	83	88	95	98
	St	30	55	67	81	87	94	98

(Continued)



<u>Number of Cores</u>		<u>+10</u>	<u>+20</u>	<u>+25</u>	<u>+35</u>	<u>+40</u>	<u>+50</u>	<u>+60</u>
20	Sm	31	58	69	85	90	96	99
	St	32	58	69	84	89	95	98
22	Sm	34	61	72	87	92	97	99
	St	34	60	71	86	91	96	99
24	Sm	35	63	74	89	93	98	99
	St	35	62	73	88	92	97	99
26	Sm	37	66	76	90	94	98	99
	St	36	64	75	90	93	98	99
28	Sm	37	67	77	91	95	98	100
	St	37	66	77	91	94	98	100
30	Sm	38	68	79	92	96	99	100
	St	38	68	79	92	95	99	100
32	Sm	40	71	81	93	96	99	100
	St	39	70	80	93	96	99	100
34	Sm	42	72	82	94	97	100	100
	St	40	72	82	94	97	99	100
36	Sm	42	72	83	95	97	99	100
	St	41	73	83	95	97	100	100
38	Sm	44	75	85	95	98	100	100
	St	42	74	84	95	98	100	100
40	Sm	44	76	85	96	98	100	100
	St	43	75	85	95	98	100	100

\* Simulation

\*\* Statistical

Table VIII. Percent accuracy of pH sampling on the Cameron farm data as determined by simulation and statistical techniques.

Mean pH value for field = 6.4

Interval ( $\pm$  pH units from the mean)

<u>Number of Cores</u>		<u><math>\pm .1</math></u>	<u><math>\pm .2</math></u>	<u><math>\pm .3</math></u>	<u><math>\pm .4</math></u>	<u><math>\pm .5</math></u>	<u><math>\pm .6</math></u>	<u><math>\pm .7</math></u>
2	* Sm	26	49	67	80	89	95	97
	** St	25	45	65	80	85	90	95

(Continued)



<u>Number of Cores</u>		<u>± .1</u>	<u>± .2</u>	<u>± .3</u>	<u>± .4</u>	<u>± .5</u>	<u>± .6</u>	<u>± .7</u>
4	Sm	35	64	83	93	98	99	100
	St	35	60	80	90	96	99	99
6	Sm	43	75	91	97	99	100	100
	St	40	70	90	96	99	100	100
8	Sm	49	80	95	99	100	100	100
	St	47	80	93	99	99	100	100
10	Sm	54	86	97	100	100	100	100
	St	52	85	95	99	100	100	100
12	Sm	59	89	98	100	100	100	100
	St	55	90	98	100	100	100	100
14	Sm	62	92	99	100	100	100	100
	St	60	92	98	100	100	100	100
16	Sm	64	94	99	100	100	100	100
	St	64	93	99	100	100	100	100
18	Sm	68	95	100	100	100	100	100
	St	67	94	99	100	100	100	100
20	Sm	70	96	100	100	100	100	100
	St	69	95	99	100	100	100	100
22	Sm	72	97	100	100	100	100	100
	St	71	96	100	100	100	100	100
24	Sm	74	98	100	100	100	100	100
	St	73	97	100	100	100	100	100
26	Sm	76	98	100	100	100	100	100
	St	75	98	100	100	100	100	100
28	Sm	78	99	100	100	100	100	100
	St	77	98	100	100	100	100	100
30	Sm	79	99	100	100	100	100	100
	St	79	99	100	100	100	100	100
32	Sm	81	99	100	100	100	100	100
	St	80	99	100	100	100	100	100
34	Sm	83	99	100	100	100	100	100
	St	82	99	100	100	100	100	100

(Continued)



<u>Number of Cores</u>		<u>±</u> .1	<u>±</u> .2	<u>±</u> .3	<u>±</u> .4	<u>±</u> .5	<u>±</u> .6	<u>±</u> .7
36	Sm	84	100	100	100	100	100	100
	St	83	99	100	100	100	100	100
38	Sm	86	100	100	100	100	100	100
	St	84	100	100	100	100	100	100
40	Sm	86	100	100	100	100	100	100
	St	85	100	100	100	100	100	100

\* Simulation

\*\* Statistical

---





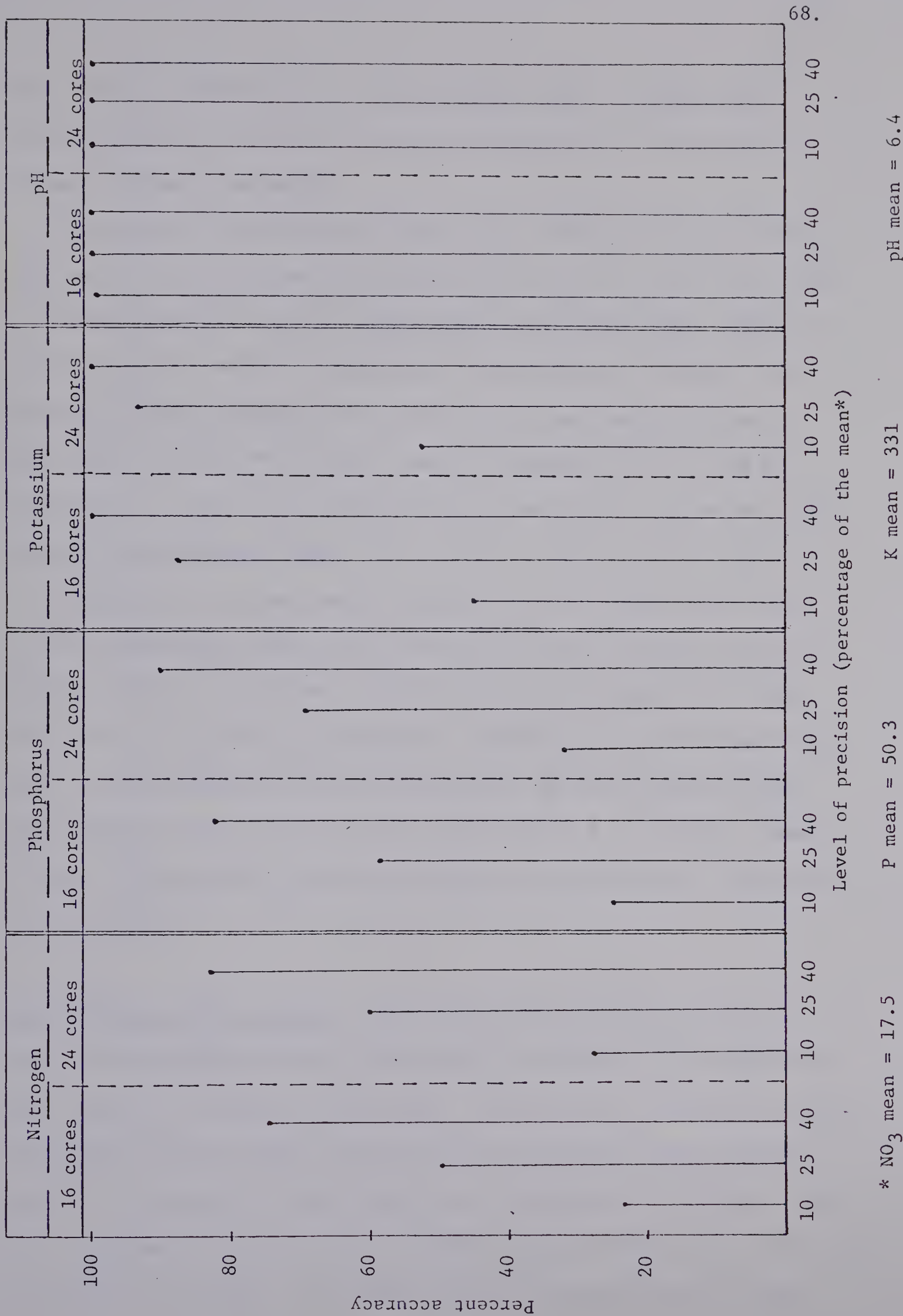
for nitrogen sampling for a level of precision of  $\pm 5$  lb/ac from the mean. Phosphorus sampling is 65% accurate using a  $\pm 7$  lb/ac interval and potassium is 93% accurate using a  $\pm 50$  lb/ac interval with 16 cores. The same number of cores with a  $\pm .5$  interval about the mean gives a 100% accuracy for pH sampling. Assuming a required 80% accuracy with the same intervals as above, one can see from the tables that the number of cores required is 10, 30, 8, and 1 for  $\text{NO}_3$ , P, K, and pH respectively. The data on the distribution of the number of cores taken shows that 80% of the samples are represented by at least 9 cores. This gives a sampling accuracy of at the very least (but most likely greater than) 78%, 51%, 84%, and 100% for  $\text{NO}_3$ , P, K, and pH, respectively, using the above levels of precision.

Figure 3 shows a comparison of the percent accuracy of sampling for each of the four Laboratory determinations. The graph compares the accuracy for 16 and 24 cores with three levels of precision expressed as a percentage of the mean. The results show that when working with a percentage of the mean that nitrogen and phosphorus are the least accurately determined nutrients. Potassium shows a considerable improvement and pH is very accurately determined with both 16 and 20 cores at the 10%, 25%, and 40% levels of precision.

Generally, the sampling results from the Cameron farm correspond to those obtained by Hemingway (op. cit.) and Holland et al. (op. cit.). When the level of precision is expressed as a percentage of the mean, nitrogen has a sampling error comparable to that of phosphorus. This may well be attributed to the lower field mean for nitrogen. Hemingway (op. cit.), in his sampling study, shows that soils with low means tend to



Figure 3. Comparison of percent accuracy of sampling nitrogen, phosphorus, potassium, and pH with level of precision expressed as a percent of the mean. Cameron farm data used.





have a higher sampling error on a percentage basis. In this case, absolute intervals rather than intervals expressed as a percentage of the mean are more reasonable.

The results of Holland et al. (op. cit.) working in apple orchards show that with 24 cores and intervals of  $\pm 9.4$  lb/ac,  $\pm 48.2$  lb/ac, and  $\pm .24$  units for P, K, and pH, respectively, that 80% accuracy (80% confidence interval) could be achieved for the sampling of orchards. The Cameron farm data indicates that with 24 cores and similar intervals that an 87%, 97% and 99% accuracy could be achieved for P, K, and pH, respectively. Part of the inconsistency in the results may be due to the different chemical tests used.

The inference of the results obtained from the Cameron farm study is that approximately 80% of the time the field mean obtained will fall within  $\pm 4$  lb/ac,  $\pm 10$  lb/ac,  $\pm 35$  lb/ac, and  $\pm .15$  units of the true mean for N, P, K, and pH, respectively. Expressed as a percentage of the mean, the sampling mean will fall within 45%, 40%, 23%, and 5% of the true population mean four out of five times for N, P, K, and pH, respectively. As yet, however, these inferences are unverified for other parts of the province.

Field size and sample numbers: The Cameron farm data also served as a basis for the determination of the effect of field size on the number of cores required to compose a soil sample. Using Stein's two-stage sampling formula with varying levels of precision and confidence limits (percent accuracy) the number of cores required was determined for different field sizes. A summary of the results obtained is shown in Tables IX, X, XI, and XII for N, P, K, and pH. The tables show the average number of cores









Table X. Average number of cores required for phosphorus sampling on different field sizes.  
Cameron farm data used. Range indicated for each average.

Field size (acres)	Number of runs	Interval (lb/ac )	Confidence Limits							
			70%		80%		90%		95%	
			Average	Range	Average	Range	Average	Range	Average	Range
.5	693	+ 5 + 8 +10	29 12 8	5-73 4-31 3-24	46 18 12	5-114 5-43 4-37	80 31 20	5-172 5-82 5-52	120 47 30	5-212 5-115 5-72
1.	684	+ 5 + 8 +10	33 13 9	5-71 4-37 3-25	51 20 13	5-125 5-51 4-38	85 34 21	10-206 5-83 5-53	124 49 31	10-227 5-160 5-75
5.	445	+ 5 + 8 +10	42 17 11	25-76 10-29 7-19	65 25 17	40-112 15-44 10-29	107 41 27	65-188 25-71 18-48	152 59 39	90-232 35-101 25-67
10.	142	+ 5 + 8 +10	44 17 11	32-63 12-26 8-16	67 26 17	49-99 18-39 12-26	109 42 27	80-157 30-62 20-41	155 60 39	110-230 45-91 28-57
15.	152	+ 5 + 8 +10	41 16 11	33-49 12-19 8-14	63 25 16	50-81 20-30 12-19	103 41 26	80-132 30-51 20-36	146 57 37	110-183 45-71 29-45
20.	114	+ 5 + 8 + 10	43 17 11	35-50 15-19 9-14	66 26 17	55-75 24-29 13-19	109 42 27	90-119 35-50 24-30	154 60 39	125-169 55-67 30-44
25.	60	+ 5 + 8 +10	42 17 11	40-45 16-20 10-13	64 25 16	60-69 23-29 15-18	106 42 27	100-116 41-45 25-29	150 59 38	140-160 55-64 35-40



Table XI. Average number of cores required for potassium sampling on different field sizes.  
Cameron farm data used. Range indicated for each average.

Field size (acres)	Number of runs	Interval  (lb/ac )	Confidence Limits									
			70%		80%		90%					
			Average	Range	Average	Range	Average	Range				
.5	693	+ 25 + 35 + 50	14 7 4	4-45 3-28 1-16	22 11 6	5-62 4-31 2-20	38 19 10	5-103 5-57 4-32	56 29 14	5-157 5-91 4-46		
		1.0	684	+ 25 + 35 + 50	15 8 4	5-42 4-22 1-11	22 12 6	5-57 4-31 3-16	38 19 10	5-83 5-56 4-27	54 28 14	5-156 5-68 5-36
				5.0	445	+ 25 + 35 + 50	17 9 5	10-30 5-15 2-10	26 13 7	10-41 5-25 4-11	42 22 11	15-73 10-36 5-20
10.0	142					+ 25 + 35 + 50	17 9 5	14-29 7-14 2-9	27 14 7	20-40 10-21 5-10	43 22 11	30-66 15-34 9-22
		15.0	152			+ 25 + 35 + 50	21 11 6	19-25 10-14 4-9	32 17 9	29-45 15-20 6-10	53 27 14	45-66 23-35 11-20
				20.0	114	+ 25 + 35 + 50	22 12 6	19-25 10-14 4-9	34 18 9	30-41 15-20 6-10	55 28 14	50-60 25-34 11-15
25.0	60					+ 25 + 35 + 50	22 11 6	20-25 10-14 4-9	33 17 9	31-35 16-19 6-10	54 28 14	51-59 25-30 11-15



Table XII. Average number of cores required for pH sampling on different field sizes.  
Cameron farm data used. Range indicated for each average.

Field size (acres)	Number of runs	Interval (pH units)	Confidence Limits							
			70%		80%		90%		95%	
			Average	Range	Average	Range	Average	Range	Average	Range
.5	693	+ .2	4	1-11	6	2-22	11	3-31	16	4-43
		+ .3	1	1-9	3	1-12	5	2-16	7	2-22
		+ .5	1	1-4	1	1-5	2	1-6	3	1-11
1.	684	+ .2	5	1-13	7	2-17	11	3-27	16	4-37
		+ .3	2	1-7	3	1-9	5	2-15	7	2-21
		+ .5	1	1-4	2	1-5	2	1-6	3	1-11
5	445	+ .2	5	2-12	8	4-16	13	5-26	18	5-37
		+ .3	3	1-5	4	1-10	6	2-12	8	2-16
		+ .5	1	1-4	2	1-4	2	1-4	3	1-7
10	142	+ .2	8	2-12	11	5-16	18	8-22	25	10-32
		+ .3	4	1-11	5	1-11	8	2-13	12	5-16
		+ .5	2	1-6	2	1-7	3	1-8	5	1-10
15	152	+ .2	6	3-10	7	6-15	14	10-20	20	14-24
		+ .3	3	1-5	4	1-5	7	4-10	9	6-15
		+ .5	1	1-4	2	1-5	3	1-5	4	2-5
20	114	+ .2	6	4-8	8	6-10	14	11-15	19	15-21
		+ .3	3	2-4	4	2-5	6	5-8	9	7-10
		+ .5	1	1-3	2	1-4	3	2-5	3	2-5
25	60	+ .2	6	4-8	8	6-10	13	11-15	19	16-20
		+ .3	3	2-4	4	3-5	6	5-8	9	7-10
		+ .5	1	1-3	2	1-3	3	2-5	4	2-5



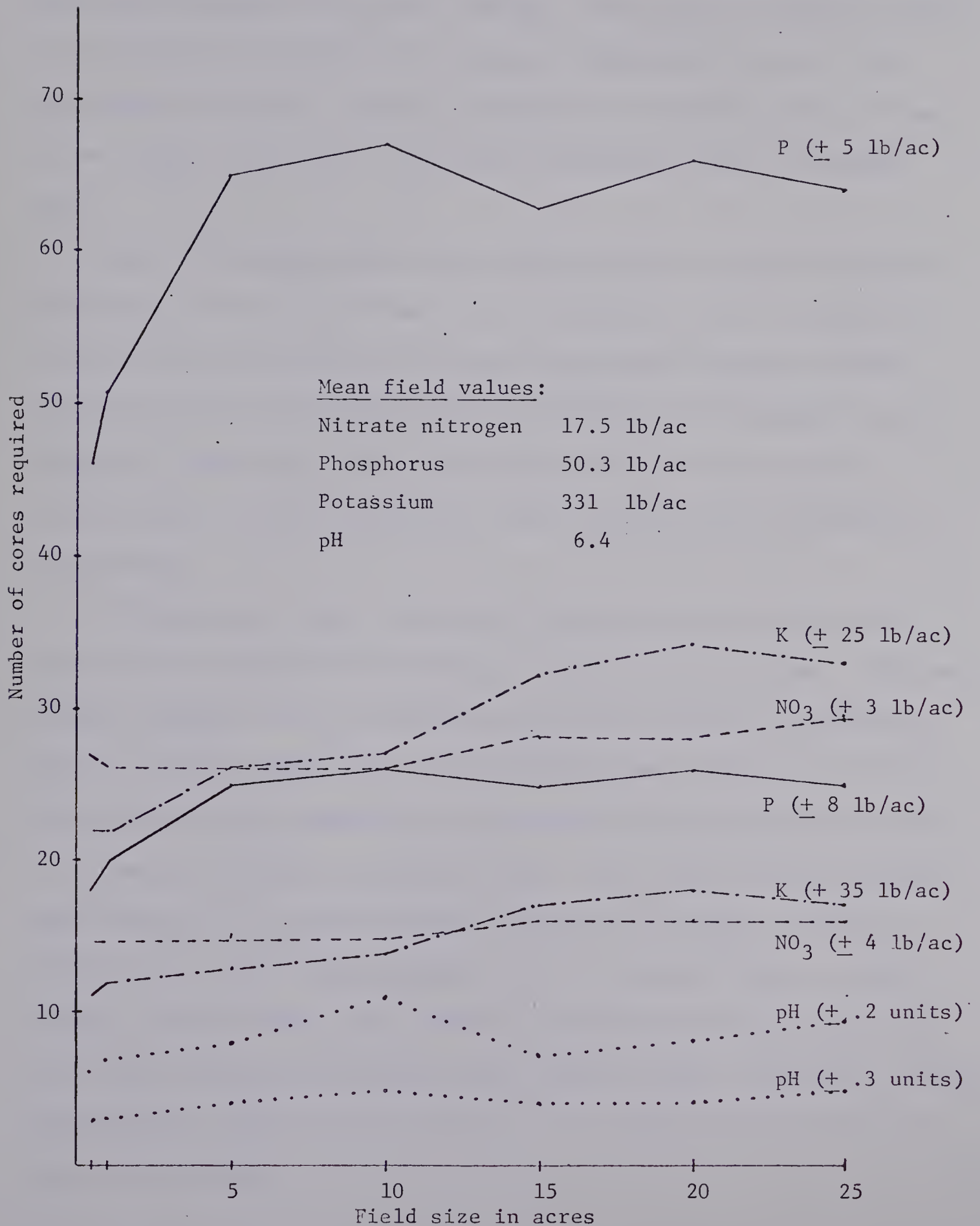
required and the range for each field size. The number of runs for each field size is simply the number of different locations within the original at which a calculation was made for the number of cores required. As the field size increased, the number of runs decreased because of the difficulty in finding new locations for field sizes approaching the size of the original field.

The results show that in all cases there is no substantial continual increase in the number of cores required (on the average) as the field size increases from .5 to 25 acres. For confidence limits of 80% and an interval of  $\pm 4$  lb/ac, 15 cores are required on the average for a .5 acre field for nitrogen sampling, while 16 are required for a 25 acre field. The increase is definitely insignificant. Phosphorus, on the other hand, for the same confidence limits and an interval of 16 lb/ac requires 18 cores on the average for a .5 acre field and 25 for a 25 acre field. The increase is not linear, however, for 5 acres also require 25 cores to meet the above specifications. Potassium behaves similarly to phosphorus with 11 cores required for .5 acres and 17 for a 25 acre field using 80% confidence limits and an interval of  $\pm 35$  lb/ac. Potassium has more of a linear increase than phosphorus, but tends to level out at 15 acres. Soil reaction follows a very similar trend to nitrogen with very little increase in variance and thus very little increase in the number of cores required. For 80% confidence limits with a precision of  $\pm .3$  pH units the number of cores required for a .5 acre field is 3 and for a 25 acre field is 4. Figure 4 shows a graphical comparison of the number of cores versus field size for N, P, K, and pH using 80% confidence limits and two levels of precision: the ones shown above and another more rigorous set.





Figure 4. Number of cores required using 80% confidence limits for nitrogen, phosphorus, potassium, and pH sampling on different field sizes.





As previously noted, the range of the number of cores required for each different field size was recorded. As the field size increases, the range decreases for the four analyses. This gives an indication that as the field size increases the variances calculated for each of the measurements stabilize. However, because of the limited number of runs on the larger field sizes, one cannot be entirely certain in stating this.

Where a wide range exists, the average number of samples required persistently falls at the lower end of the range. This is a good indication that the mode will be lower than the average and that a skewed distribution most likely exists for the population of variances usually obtained. Greater than 50% of the variances will be such that the actual number of cores required will tend to be less than indicated by the average.

Hemingway (op. cit.) states that field size does not affect the number of cores required for sampling a field for any nutrient. Peck and Melsted (op.cit.), on the other hand, state that the work of Keogh and Maples on alluvial soils show that 20 borings were needed for a 20 acre field, while 15 were needed for 10 acres and 10 were needed for 5 acres to achieve an arbitrary, acceptable limit. The results from the Cameron farm data point out that the major increase in the number of cores required, if any, takes place between .5 and 5 acres and tends to remain constant beyond 5 acres. Since 99% of the fields generally sampled for soil tests in Alberta are greater than 5 acres in size, the number of cores composing the sample is more important in the determination of accuracy than the field size.



Sampling on a township and larger area basis: Since no literature exists on sampling for N,P,K, and pH on an area larger than field size and since this data is useful for determining the accuracy of mapping, computer programs were written to obtain this information using the farm soil sample data. In using the data it should be recalled that farm soil samples are usually not individual cores, but rather samples composed of individual cores. Thus, the results cannot theoretically be compared to the sampling on a field basis since one is working with the means of the individual cores, not the individual cores, themselves.

Stein's two-stage sampling formula was used and predictions were made for the number of samples required for both a township and a critical circle composed of 37 townships. A diagram of the critical circle used may be found in Appendix IIb. The results for the actual distribution of sample numbers are shown in Tables XIII and XIV, respectively.

Table XIII presents the results obtained for the average number of samples required within a township and a circle as well as the minimum number of samples required that will ensure that at least 80% of these areas are adequately represented. The calculations were carried out for each of the four nutrients with selected crop groupings and time periods associated with the mapping of each nutrient using an 80% confidence interval with a fixed level of precision for each nutrient. At least 6 samples within a circle and 4 within a township had to be present before a calculation was performed, ensuring samples from two different farms on the average.



Table XIII. Average number of samples required within a township and critical circle for nitrogen, phosphorus, potassium, and pH on different crop groupings and years for 80% sampling accuracy.

Years	Nutrient - - - Crop	Nitrogen on fallow		Nitrogen on stubble & grass-legume crops		Phosphorus on fallow, stubble & grass-legume crops		Potassium on all crops		pH on all crops	
		Ave.	* 80%	(+ 7.5 lb/ac) Sept.-April	(+ 5 lb/ac) Sept.-April	Ave.	80%	(+ 50 lb/ac) All months	Ave.	(+ .3 units) All months	80%
	Interval - - - Months - - -										
	Area size										
62-69	Circle	17	22		18	25	37	-	6	8	
65-69	Circle	17	21		17	24	37	24	6	7	
62-69	Twp.	19	27		19	25	32	-	5	7	
65-59	Twp.	18	25		18	25	31	22	5	6	
68-69	Circle	10	16		11	15	20	26	6	8	
67-68	Circle	18	26		22	22	35	24	5	7	
66-67	Circle	16	22		18	25	38	21	6	7	
65-66	Circle	12	15		16	30	44	22	7	8	
68-69	Twp.	8	11		12	14	16	22	5	7	
67-68	Twp.	20	32		21	25	31	23	5	6	
66-67	Twp.	14	23		20	24	32	20	5	6	
65-66	Twp.	12	19		17	29	32	22	5	7	
** 62-69	Circle	39	49		49	56	83	-	3	3	
** 65-69	Circle	37	47		46	54	82	93	3	3	
*** 62-69	Circle	13	16		17	28	41	-	6	7	
*** 65-69	Circle	12	15		16	27	41	23	6	7	

\* 80% of the places sampled require less than or equal to that number of samples.

\*\* Interval changed -- N on fallow to + 5, N on stubble, etc. to + 3, P to + 5, K to + 25, and pH to + .5

\*\*\* Crop changed -- N on all crops, N on stubble only, P on all crops, K and pH on fallow and stubble only.





Table XIV. Average number of samples within a township and critical circle for different crop groups and months and years.

Years	Months	*Crops	No. of samples within a critical circle		No. of samples within a township		No. of townships (with samples) within a critical circle	
			Ave.	**80%	Ave.	**80%	Ave.	**80%
62-69	all	all	585	15	30	5	21	3
62-69	all	F, S, LG	599	34	29	6	22	5
62-69	Sept.-Apr.	S, LG	424	26	21	4	21	5
62-69	Sept.-Apr.	F	185	13	9	2	20	5
65-69	all	all	463	15	24	5	21	3
68-69	all	all	125	14	8	2	16	4
67-68	all	all	165	13	9	2	18	3
66-67	all	all	162	16	10	2	17	4
65-66	all	all	80	9	6	2	14	3
68-69	all	F, S, LG	121	14	8	2	16	4
67-68	all	F, S, LG	168	19	9	2	19	5
66-67	all	F, S, LG	157	16	10	2	17	5
65-66	all	F, S, LG	78	10	6	2	14	3
68-69	Sept.-Apr.	S, LG	88	11	6	2	15	4
67-68	Sept.-Apr.	S, LG	120	14	7	2	16	5
66-67	Sept.-Apr.	S, LG	112	13	8	2	16	4
65-66	Sept.-Apr.	S, LG	54	7	5	1	12	3
68-69	Sept.-Apr.	F	40	6	3	1	12	3
67-68	Sept.-Apr.	F	53	6	4	1	14	4
66-67	Sept.-Apr.	F	48	6	4	1	13	3
65-66	Sept.-Apr.	F	26	4	3	1	10	2

\* F = fallow, S = Stubble, LG = legume-grass

\*\* 80% of the places sampled have greater than or equal to that number of samples



A comparison of the results for the circle and the township in the same time period shows that there is very little difference in the average number of samples required, despite the size difference. This re-emphasizes the fact that the number of samples, and not the size of the area is more important in obtaining the required accuracy. The number required that would represent at least 80% of the area sampled (with 80% confidence at a given level of precision) is more variable. A comparison of the years 1962 to 1969 for both the township and the circle shows that the nitrogen 80% sample requirements tend to be higher in the township than in the circle, indicating more variation on a local basis than over a whole area. Phosphorus and particularly potassium and pH show an almost opposite trend, indicating more variation as area size increases.

A year to year comparison shows that the number of samples required changes, especially for nitrogen and phosphorus. This strongly suggests that the mean levels of these nutrients fluctuate annually (assuming the variance increases as the mean does). For example, the fall of 1967 appears to have a fairly high level of nitrate accumulation, while the fall of both 1968 and 1965 appears to have a smaller accumulation. The fall of 1965 appears to have a very high level of phosphorus, while the fall of 1968 tends to have a lower content. Potassium and pH do not appear to vary extensively from year to year.

The number of samples within the critical circle as well as the largest number of samples that is available in at least 80% of the sampling area with varying time periods and crops is shown in Table XIV. In addition, the table shows results for the number of samples within a township, and the number of townships (with samples) within the critical circle.



Table XIV shows several important aspects related to sampling density of the data by years and crops. There is an obvious and reasonable decrease in the average number of samples within any area when taking into account all the years as compared to a few years or one year. However, this same drop is not as prominent when considering the number of townships represented by samples. Taking into account the data from all years, months, and crops, an average of 21 townships is represented in the critical circle; the same critical circle for 1967-68 has an average of 18 townships represented. However, each township is represented by fewer samples (30 compared to 9). The results would tend to indicate, then, that the relative density of represented townships per critical circle is very similar for any reasonable time period, but sample numbers will change considerably.

Selection of particular crop types also changes the number of samples available, but as before there is not as noticeable a change in the number of represented townships per critical circle. In fact, for the sampling year of 1967-68, taking into account all crops, an average of 18 townships with samples is available in the critical circle; whereas, just taking into account fallow for the same sampling year, 14 townships are represented. At the same time the actual sample number decreases from 165 to 53 in the critical circle, and from 9 to 4 in a township. Looking at the summary over the years of 1962-69, one finds an unexpected increase in the average number of samples within a critical circle after samples from vegetable crops and breaking have been excluded. Such an increase corresponding to a loss of samples from certain crops gives an excellent indication that these crops are from fringe areas where samples are very few and far between. This, in effect, will lower the average in





the critical circle. It should be mentioned that most likely breaking and not the samples from vegetables are from the fringe areas.

Some years are more heavily sampled than others. The sampling year of 1967-68 tends to give the largest average of samples within the critical circle for any combination of crops. On the other hand, the sampling year of 1966-67 appears to be the heaviest on a township basis. The latter trend is an indication of a heavier concentration of samples within a township, but at the same time a relative spread of the individual townships.

The main purpose of this particular study was to determine the number of samples required within a critical circle for a certain degree of accuracy and level of precision; and to relate this to the number of samples available within the critical circle. Such data is important in that it gives an indication as to the accuracy of the mapping function and, in addition, it can be used to determine the type of control and elimination needed within the critical circle of the mapping function.

A comparison of the two tables allows one to obtain an idea of the accuracy of mapping that can be achieved. For example, considering nitrogen on fallow in the sampling year 1967-68, there is a requirement of at least 26 samples to be accurate on 80% of the mapping area using a circle. Table XIV shows that for the same time period and cultural practice that on the average 53 samples exist, but 20% of the area covered has less than 6 samples where 26 are required. Nitrogen on stubble and legume-grass shows a requirement of 29 samples within the critical circle to be 80% accurate for 80% of the area in the sampling year 1967-68. Table XIV shows an average of 120 samples present with at least 14 in 80% of the area. Phosphorus over all the years and on all crops except





vegetables and breaking shows a requirement of 37 samples four out of five times within the critical circle. The results show that for 80% of the area covered, the circle contains at least 34 samples with an average of 599 samples. Potassium, on all crops and months from 1965 to 1969, shows a requirement of at least 36 samples, and the results show an average of 463 with 80% of the area covered having at least 15. Soil pH is the least variable of the four analyses with a maximum of 8 samples required to ensure 80% coverage.

The results indicate, except in the case of pH, that 80% accuracy for 80% of the areas where samples are available is not always obtainable. This is especially true where fewer sample numbers exist and where only data from one year or only a few crop types are used. An exclusion of samples from new breaking, appears to eliminate a large area relative to the number of samples, thus increasing the average number of samples usually available within the critical circle. The fringe areas and areas where very few samples exist will be the least accurately mapped areas.



### Summaries by Soil Area and County

Preliminary: A prelude to any intensive summary is a brief, general look at the data and any factors that might influence it. Such a study allows one to formulate rational opinions on how best to arrange and summarize such data. In addition, it should reveal some of the general trends in the data and make one aware of any bias that might be latent in the data.

In the previous section a discussion of the number of cores taken, the field size, and the average number of acres represented by one core gives one an idea as to the type of sampling generally carried out in the province and the accuracy of the sampling on the average farm. Further analysis of the data shows that an average of 3 samples is usually sent in by one farmer. Thus, although approximately 15,000 samples may have been received in the sampling year of September, 1967 to April, 1968, these probably represent only 5,000 farms. Tables XIII and XIV, which are listed in the previous section, give an idea of the number of samples per township and over a group of townships.

A preliminary look at some of the results showing the nutrient availability of Alberta soils on stubble and fallow crops over different years is shown in Table XV. An examination of the table points out one of the first noticeable characteristics of the A.S.F.T.L. data - - nitrogen and phosphorus tend to have very large standard deviations, frequently approaching the value of the mean. The coefficients of variation are approximately .68, 1.14, .74, and .81 for nitrogen on fallow and stubble and phosphorus on fallow and stubble, respectively. On the other hand,



Table XV. Summary of the nutrient availability of Alberta soils under fallow and stubble crops over different years. Months: Sept.-Apr.

Sampling year		Nitrogen		Phosphorus		Potassium		pH	
		F	St	F	St	F	St	F	St
62-63	Mean	34	16	40	35	-	-	6.6	6.7
	S.D.	20	15	31	32	-	-	.7	.6
	No.	972	1777	972	1777	-	-	972	1777
63-64	Mean	47	22	38	37	-	-	6.6	6.7
	S.D.	26	16	28	32	-	-	.7	.7
	No.	1555	2690	1555	2690	-	-	1555	2690
64-65	Mean	23	14	43	40	-	-	6.6	6.7
	S.D.	15	14	34	31	-	-	.7	.7
	No.	994	2197	994	2197	-	-	994	2197
65-66	Mean	35	15	46	43	606	517	6.5	6.6
	S.D.	21	14	29	32	264	233	.7	.7
	No.	1858	3596	1858	3596	1858	3596	1858	3596
66-67	Mean	35	11	39	35	613	503	6.5	6.6
	S.D.	22	14	26	28	265	242	.6	.6
	No.	3342	7445	3342	7445	3342	7445	3342	7445
67-68	Mean	41	17	38	39	628	518	6.5	6.5
	S.D.	23	17	29	28	280	245	.7	.6
	No.	4012	8945	4012	8945	4012	8945	4012	8945
68-69	Mean	20	8	28	27	664	567	6.8	6.8
	S.D.	18	11	21	23	304	277	.7	.7
	No.	2799	5833	2799	5833	2799	5833	2799	5833
62-69	Mean	34	14	38	36	-	-	6.6	6.6
	S.D.	23	16	28	29	-	-	.8	.7
	No.	15094	32127	15094	32127	-	-	15094	32127
64-69	Mean	33	13	37	36	622	520	6.5	6.6
	S.D.	23	15	28	29	280	249	.7	.8
	No.	11799	25752	11799	25752	11799	25752	11799	25752



those for potassium and pH are substantially less variable with coefficients of variation usually around .45, .48, .12, and .11 for potassium on fallow and stubble and pH likewise, in that order. The table also shows that there are definite differences in levels of nitrogen on stubble and fallow and, similarly, for potassium on stubble and fallow, but the levels of phosphorus and pH do not appear to be significantly affected by any particular cultural practice. However, as the results in the table point out, phosphorus and especially nitrogen vary considerably from year to year. Thus, there appear to be trends in the levels of the various nutrients, and these trends are affected by crop type and year for some of the nutrients.

Table XV gives, however, a slightly inaccurate picture of the province. By referring to Table XVI, one can begin to see the reason why. Table XVI gives the results for the percentage distribution of crops within each of the 10 soil-climatic areas, as well as the percentage distribution of the number of samples by area, and similarly by crop type for all the A.S.F.T.L. data from 1962 to the present. The results show that samples from areas 5,6,7,8, and 9 (the central and northern areas) make up 73% of the samples received. Area 6, alone, contributes 28% of the total received. Thus, the first hint of bias is the inability of the samples to represent the province proportionately. The results are biased towards the areas which send in the majority of the samples. In addition, the southern areas tend to have more samples from summerfallow fields on a percentage basis, than do the more central and northern areas. This is possibly a contributing factor to the approximately 100 lb/ac difference in potassium on stubble and fallow, since potassium values tend





Table XVI. Percentage distribution of crops by area. Time period:  
All months, 1962-69.

Crop *	Areas										Crop Totals
	1	2	3	4	5	6	7	8	9	10	
<u>S.F.</u>	37	46	54	27	36	15	28	14	22	10	27
<u>Stubble</u>	53	49	41	57	52	67	58	55	53	53	56
Wheat *	28	23	23	16	23	12	26	10	17	18	18
Oats	3	2	4	7	7	13	12	12	4	7	9
Barley	12	11	8	23	16	35	15	25	22	18	21
Flax	3	6	1	2	<1	<1	<1	<1	2	6	1
Rape	2	1	2	2	3	3	3	3	10	0	3
<u>Gr-leg</u>	5	3	3	14	8	13	10	23	9	20	12
Gr-seed	<1	<1	<1	<1	<1	<1	<1	<1	3	<1	<1
Hay-pas (25)	3	2	2	6	5	6	5	12	2	11	6
Gr-leg ( 25)	2	1	1	8	3	7	5	11	7	9	6
<u>Brk-veg</u>	5	2	2	2	4	5	3	8	8	17	5
Vegetables	2	1	0	0	0	1	0	0	0	15	1
Brking	3	1	2	2	4	4	3	8	8	2	4
Area Totals	4	6	10	3	10	28	15	7	13	4	100

\* Crop means the cultural practice that preceded the sampling, eg. wheat means that a wheat crop was grown on the area previous to sampling. All underlined crops are large grouping of individual crops.



to be quite high in the southern areas.

The number of samples by crop and by area probably varies considerably from year to year. The results do, however, indicate that the majority of the samples received come from stubble fields (56%) with the majority of these from either wheat (18%) or barley (21%). Samples from fallow fields comprise 27%, all grass-legume crops\* comprise 12%, and breaking and vegetables comprise 5%. Except for Areas 1, 2, and 10, at least 75% of the samples from breaking and vegetable crops are from breaking. No two areas have the same crop distribution.

In summarizing the data over the years and over the months within the years, bias is also introduced in that the later years have contributed the majority of the samples (see Table XVII). Furthermore, certain months such as September, October, and November are the heaviest sampling months (see Table XVIII). The word "bias" is perhaps out of place in this situation, but the final results are nevertheless slanted heavily towards particular time periods. For example, 67% of the samples received, have been sampled and received since January, 1966. The samples received before that time dating back to April, 1962 account for 33% of the total. Also, 83% of the samples received have been sampled in September, October, and November, while March, April, and May account for only 9% of the samples received.

The data in Table XVIII clearly indicate that any summary over time including samples from the whole year around or just spring and fall, will definitely be weighted by values from the fall received samples. Thus, even if a distinct difference in soil test levels between seasons does exist, this difference will hardly affect the average.

\* Grass-legume crops or legume-grass crops in this text refer to any crops composed mainly of grasses or legumes or both.



Table XVII. Distribution of soil samples received by the A.S.F.T.L. based on year sampled. All months.

<u>Year</u>	<u>No. of Samples Received</u>	<u>Percent of Total</u>
1962	3460	6
1963	4960	8
1964	5410	9
1965	6760	11
1966	13850	22
1967	15490	25
1968	12500	20

Table XVIII. Distribution of soil samples received by the A.S.F.T.L. based on the month sampled. Time period 1962-69.

<u>Month</u>	<u>No. of Samples Received</u>	<u>Percent of Total</u>
Jan.	860	1
Feb.	710	1
March	1320	2
April	2720	4
May	1600	3
June	380	1
July	460	1
Aug.	660	1
Sept.	4330	7
Oct.	30100	48
Nov.	17620	28
Dec.	1730	3



The available literature indicates that most of the time there is no difference between fall and spring nutrient levels for phosphorus, potassium, and pH, and sometimes nitrogen. Filinger (op. cit.) states that of these four determinations, nitrogen is the most variable with time. Discrete differences from fall to spring have been noted occasionally. To investigate this more thoroughly, a paired t-test was run on some of the A.S.F.T.L. Nitrate Survey data (fall 1966, spring 1967) to compare fall and spring nutrient levels on 26 different fields, two from each of 13 farms across Alberta. The results show no significant difference between the fall and spring sample means for phosphorus, potassium, and pH at a 95% confidence interval. The result for nitrate nitrogen, contrary to the above, shows a highly significant difference, with the spring samples about 8 lb/ac above the fall samples.

Nitrate nitrogen in the soil is subject to extreme and often rapid fluctuations with time, especially where variable weather conditions exist. Thus, within any time period that may be selected for a study of the nutrient under field conditions, significant changes will likely occur. With due respect to this fact and since three out of four nutrients in the above statistical analysis show no significant differences due to season, it was decided that these summaries would be extended over both the fall and spring months. Thus, one sampling year usually includes all the samples that have been taken from the fall of one year to the spring of the following year. This can, of course, be varied and has been in some of the summaries.

The A.S.F.T.L. Nitrate Survey data (1966) also served as a preliminary comparison of mean values on stubble and fallow. Sample test values from adjacent stubble and fallow fields on 45 different farms throughout Alberta





were statistically analyzed using Cochran's test for equality for variances (Guenther (op. cit.) and an appropriate t-test to determine if differences exist in nutrient levels between the two practices. In the case of nitrate-nitrogen both a significant difference between variances as well as means are indicated at the .05 level of significance. Fallow has the larger mean. Phosphorus and pH show no difference in variance and no difference between means for either crop at a .05 level of significance. In the case of potassium, since 40 out of the 90 samples contained "ceiling values"\*, a special analysis had to be carried out whereby predictions based on soil area were made, replacing the "ceiling value" in the statistical calculation. The result shows no difference between variances and means for the two different cultural practices at a .05 level of significance. Thus, except for nitrate nitrogen, the study shows that usually stubble and fallow are similar in nutrient levels. Indeed, this further points out the chance of bias in Table XV where potassium levels on fallow are 100 lb/ac greater than stubble.

In order to obtain an idea of the yearly variation of soil nutrients, where no change occurred in fertilizer or cropping practice, a preliminary study was conducted using part of the Economic Productivity data in the Edmonton and Lacombe area. Except for sites 1 and 21, only the results for 1966 and 1967 were compared using Duncan's multiple range test. The results are shown in Table XIX for each of the sites.

\* "ceiling values" refer to potassium values that are greater than 600 lb/ac, but recorded as 600 + lb/ac in A.S.F.T.L. records.



Table XIX. Results of yearly variation of some selected treatments on the Economic Productivity sites.

<u>Site</u>	<u>Nitrate</u>	<u>Phosphorus</u>	<u>Potassium</u>	<u>pH</u>	<u>Crop</u>	<u>Years Compared</u>
26	* NS	NS	S	S	Forage	66-67
22	** S	NS	NS	NS	Forage	66-67
6	NS	NS	S	NS	Forage	66-67
2	S	NS	NS	NS	Forage	66-67
25	NS	S	NS	NS	Barley	66-67
23	NS	S	NS	S	Barley	66-67
21	NS	S	NS	S	Barley	65-66-67
5	NS	S	NS	NS	Barley	66-67
3	NS	NS	NS	NS	Barley	66-67
1	NS	NS	NS	NS	Barley	65-66-67

\* NS no significant difference at a 95% confidence interval

\*\* S significant difference at a 95% confidence interval

---



The results show that nitrate levels on stubble crops tend not to change significantly. Forage crops, on the other hand, show some change between years depending upon the location. Phosphorus levels tend to remain constant on forage crops but do show variation on stubble crops. Potassium levels show no difference on stubble crops, but tend to vary on forage crops. Except in a few isolated cases, pH generally shows no significant change between years. The results point out two major generalities: 1. nutrient levels under different crop types will not necessarily change similarly from year to year, and, 2. location tends to influence the changes from year to year.

Soil areas: Summaries for the 10 soil climatic areas as designated by the Soil Testing Laboratory and illustrated in Figure 1a were carried out for each of the 7 sampling years. The results from each of the years include only those samples with less than 20 acres of a field represented by one core (approximately 97% of the samples). The data that have been summed over the years include only those farm soil samples with less than 10 acres represented by one core (approximately 91% of the samples). The data summed over the years included all months.

The following discussion will deal with some of the results obtained from area summaries for nitrogen, phosphorus, potassium, and pH, respectively, for the top six inches of soil. The discussion will center about the differences and similarities in nutrient levels as affected by crop types, soil areas, and time periods. In all cases Areas 2, 6, and 9 are selected to illustrate the crop differences. These areas represent the southern, central, and northern parts of the province and at the same time have a larger number of samples relative to similar areas nearby. A



presentation of all 10 areas in this manner would be repetitious. Where conflicts exist the results from the other areas will be referred to. Data for the areas omitted are available in the Department of Soil Science, University of Alberta and at the Alberta Soil and Feed Testing Laboratory.

(a) Nitrogen - - A comparison of the nitrate nitrogen averages and standard deviations on different crop types for Areas 2, 6, and 9, is shown in Figures 5, 6, and 7, respectively. This graphical presentation allows one to grasp at a glance the difference in means and the large standard deviations that exist in such data. The results show fairly high nitrate nitrogen levels on fallow with much lower values on stubble and grass-legume crops. Straight grass crops such as grass seed show the lowest levels of nitrogen. Grass-legume crops with greater than 25% legume tend to be higher in nitrates than hay and pasture crops (under or equal to 25% legume). Vegetable crops have exceedingly high levels of nitrogen, in most cases exceeding values for fallow. Area 9 shows an unusually high value for the nitrogen level where vegetable crops have been grown. The value might well be ignored since only 19 samples were available to obtain it. The high levels of nitrate nitrogen associated with vegetable crops can most likely be attributed to the high rates of fertilizers accompanying the production of market garden crops. Land which is just broken tends to have a higher level of nitrogen than cropped land, often approaching the same level as fallow.

The results for the mean levels of nitrogen on various crops are very similar to those recorded by Wyatt et al. (op. cit.) and Newton (op. cit.), both working with Alberta soils. The only difference that can be shown is the nitrate level under vegetable crops which they found to have





Figure 5. Comparison of nitrate nitrogen averages and standard deviations on different crop types for Area 2.  
Time period: All months, 1962-68.

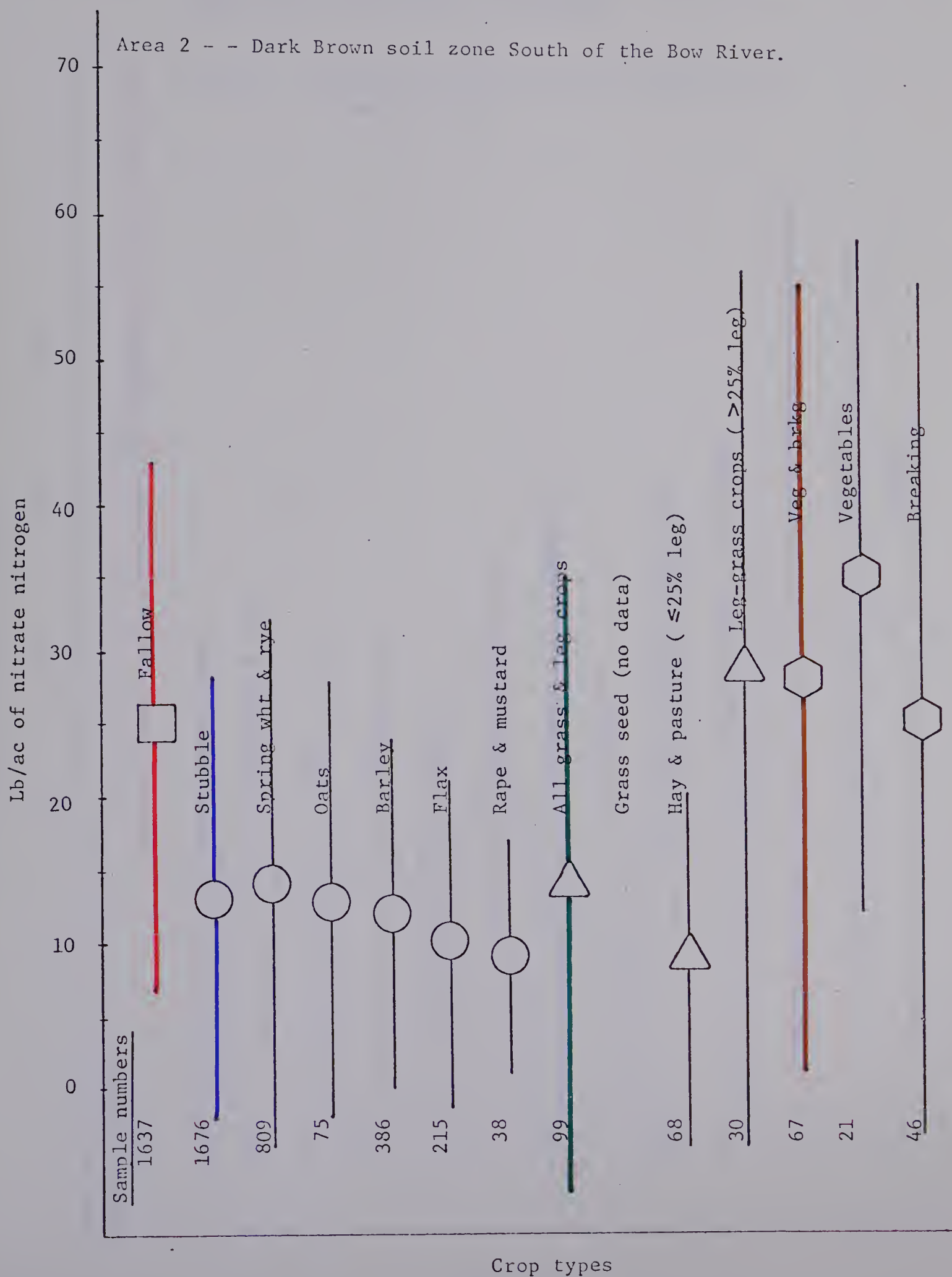




Figure 6. Comparison of nitrate nitrogen averages and standard deviations on different crop types for Area 6.  
Time period: All months, 1962-68.

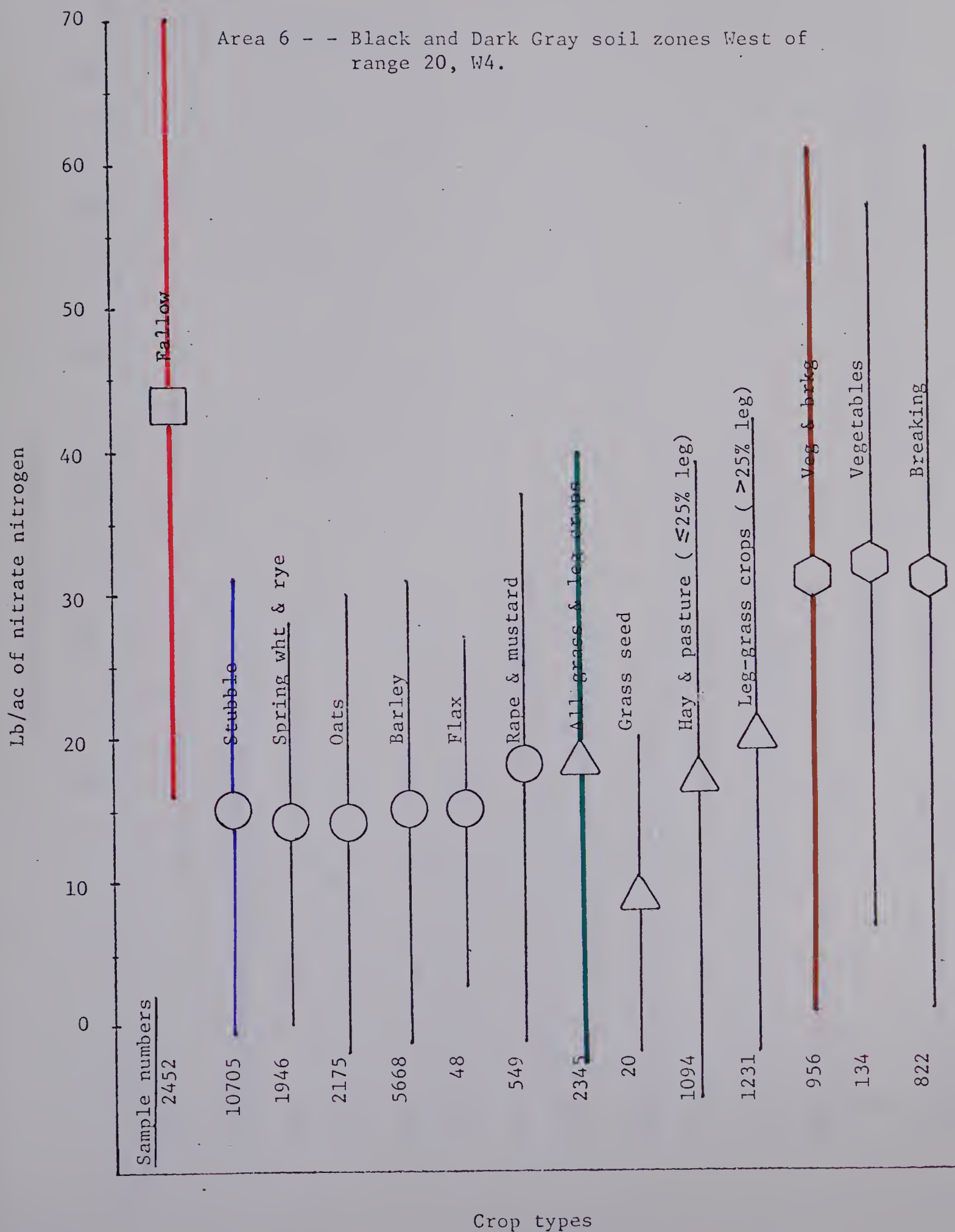
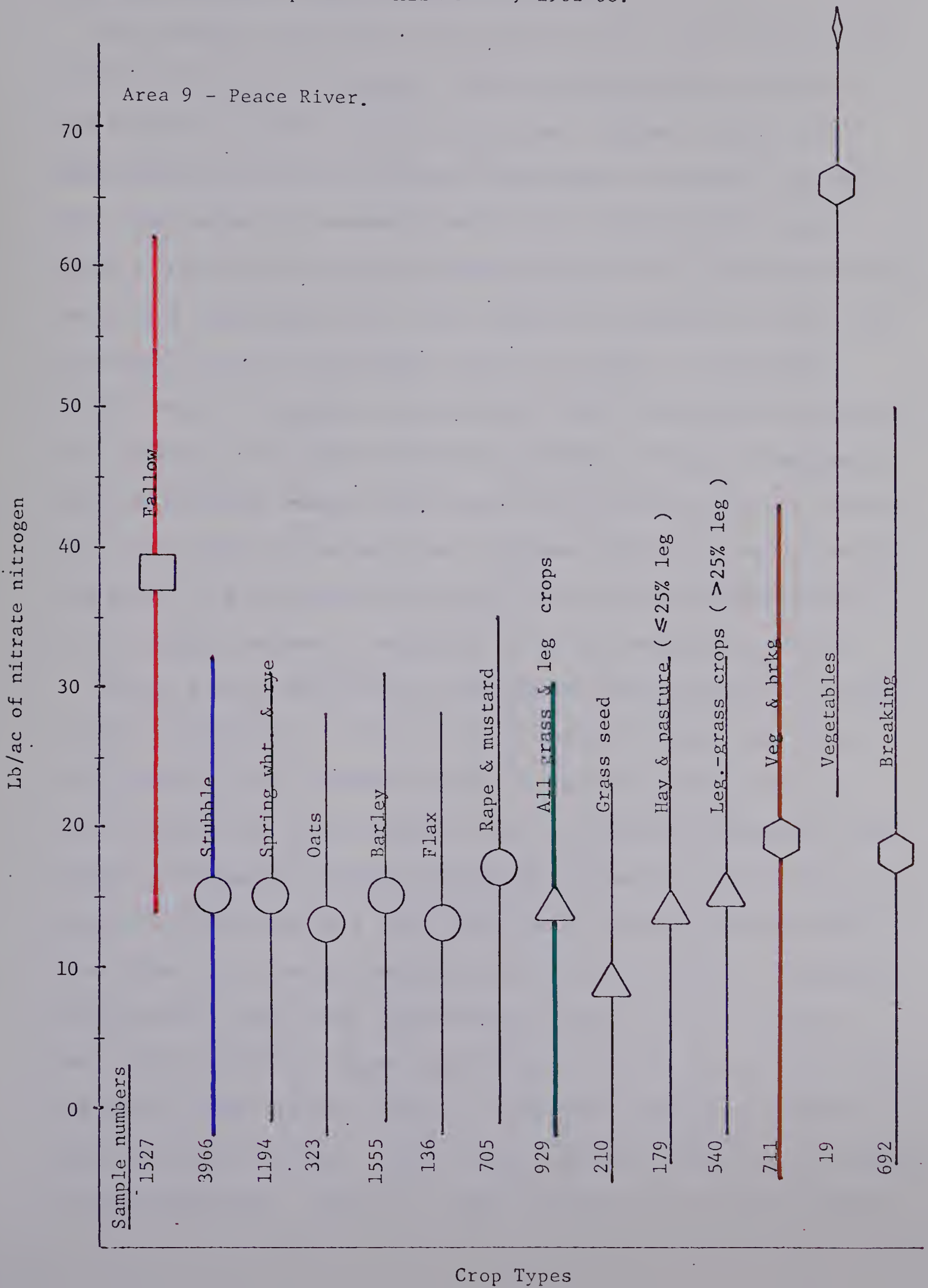




Figure 7. Comparison of nitrate nitrogen averages and standard deviations on different crop types for Area 9.  
Time period: All months, 1962-68.





intermediate values, between stubble and fallow.

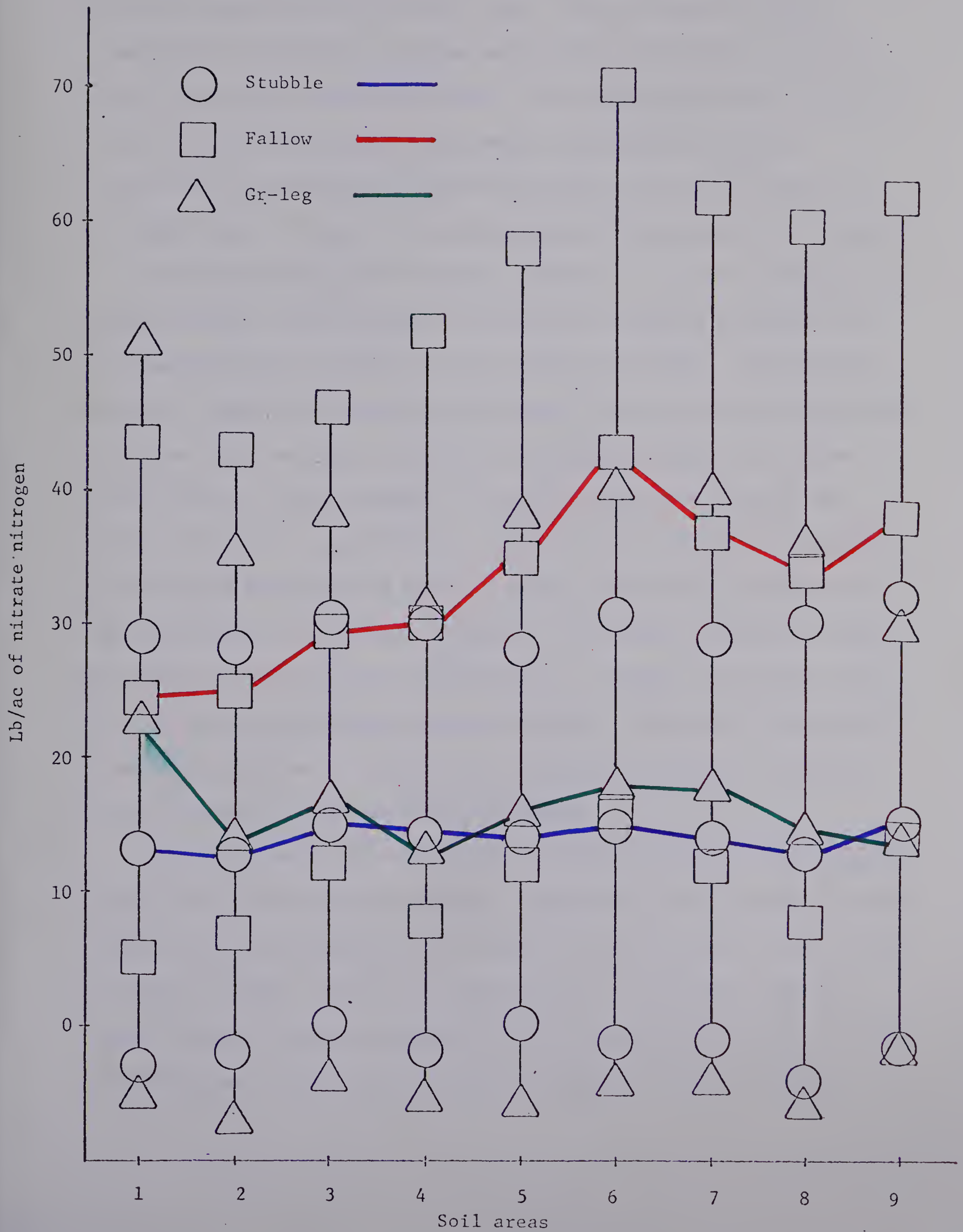
The standard deviations on all crops are quite large indicating considerable variability of results. Breaking and vegetables and fallow tend to have the largest standard deviations. However, stubble and legume-grass crops have the largest coefficients of variation (because of their lower means). The results would tend to indicate that mapping should be carried out on fallow separate from the other cultural practices. Stubble and legume-grass crops could probably be grouped since the nutrient levels under both crop groups tend to be similar. In this case, stubble crops will dominate, accounting for 82% of the samples in the long run. Because of the great variability apparently existing in vegetable crops and breaking, samples from these fields are better ignored. Even if used, their effect on the map value would most likely not even be noticeable except in a few isolated areas where very few other samples exist. On the average they make up only about 5% of the samples encountered.

Figure 8 shows the means and standard deviations of nitrate nitrogen for three crop groups in nine of the soil-climatic areas. Area 10 is not included in these summaries because it is not an exact area, but rather includes any samples from all over the province on irrigated land, although the majority of these samples come from Areas 1 and 2. The results illustrated in this manner affirm the grouping of stubble with grass-legume crops and the separation of fallow for purposes of mapping. The graph shows very little difference from area to area for nitrate levels under stubble and legume-grass crops. Areas 5, 6, and 7 where the Thin Black, Black, and Dark Gray soil zones exist tend to be slightly higher in nitrate nitrogen. Fallow shows a definite difference in nitrogen levels between areas. There is a steady increase in the nitrate level as





Figure 8. Nitrate nitrogen averages and standard deviations by crop grouping and soil area. Time period: All months, 1962-68.





one goes from the southern dryer areas with lighter coloured soils to the more central areas with darker soils. Area 6 shows the highest accumulation of nitrates on fallow with a mean of 43 lb/ac and Area 1 is the lowest with a mean of 24 lb/ac. The deviations in all areas tend to be directly proportional to the means, indicating a constant coefficient of variation for fallow from area to area. The reason for the high level of nitrate on the legume-grass crops in Area 1 is unknown.

An illustration of the variation from year to year for nitrogen is shown in Figure 9 which compares the means and standard deviations for each sampling year in Area 6 for both stubble and fallow. The results from this graph show substantial differences from year to year especially on fallow. The peak year for nitrate accumulation appears to be the fall of 1963 in which averages of 58 and 24 lb/ac were recorded for fallow and stubble, respectively. The lowest year was the fall of 1964 when the mean dropped to 28 lb/ac on fallow and the fall of 1968 when the mean dropped to 8 lb/ac on stubble. Both stubble and fallow fields for this area tend to react simultaneously to changes from year to year, with fallow showing the more pronounced effect. The graph shows that no apparent differences in variation (as indicated by standard deviation) exist from year to year on either crop type.

Figures 10 and 11 show mean nitrate nitrogen levels by soil area and sampling year for fallow and stubble, respectively. Both graphs show that there is more of a trend for soil nitrate to vary from year to year than from area to area. The fall of 1968 appears to be the year with the lowest values of nitrate nitrogen occurring in most areas, while the fall of 1963 appears to be a year in which the highest values of nitrate



Figure 9. Yearly variation of nitrate nitrogen on stubble and fallow in Area 6. Months: Sept.-Apr.

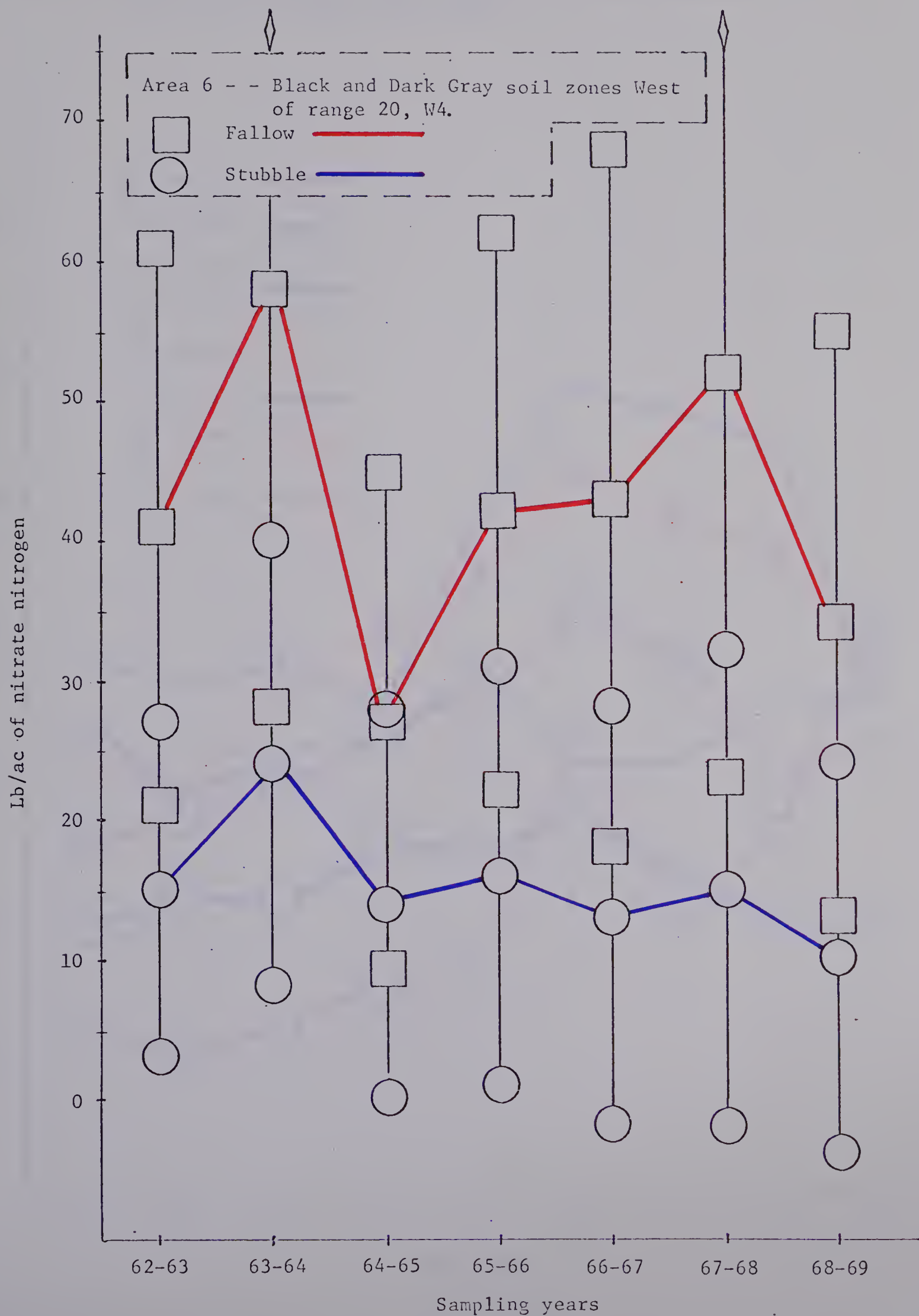




Figure 10. Nitrate nitrogen averages on fallow by sampling year and soil area. Months: Sept.-Apr.

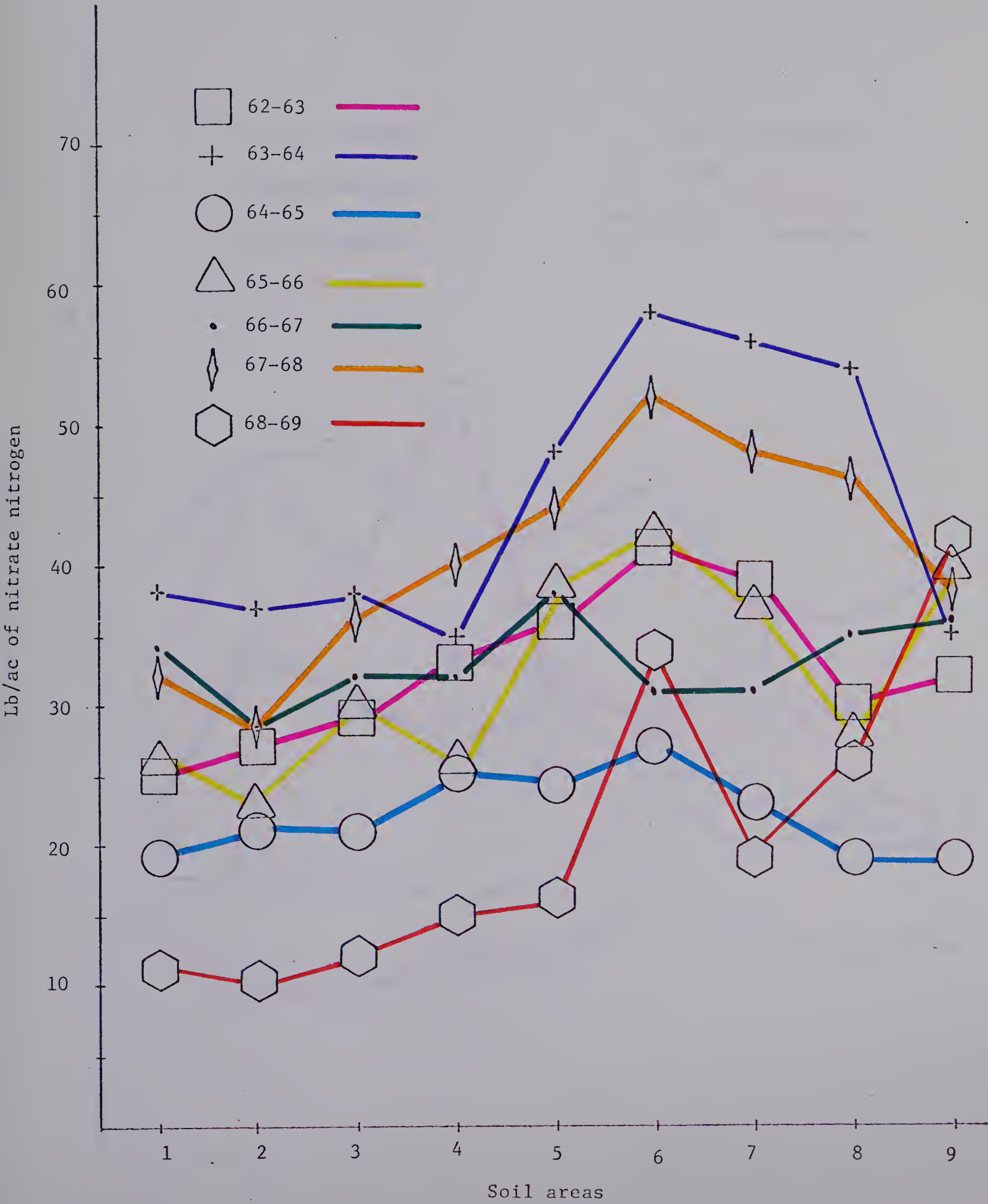
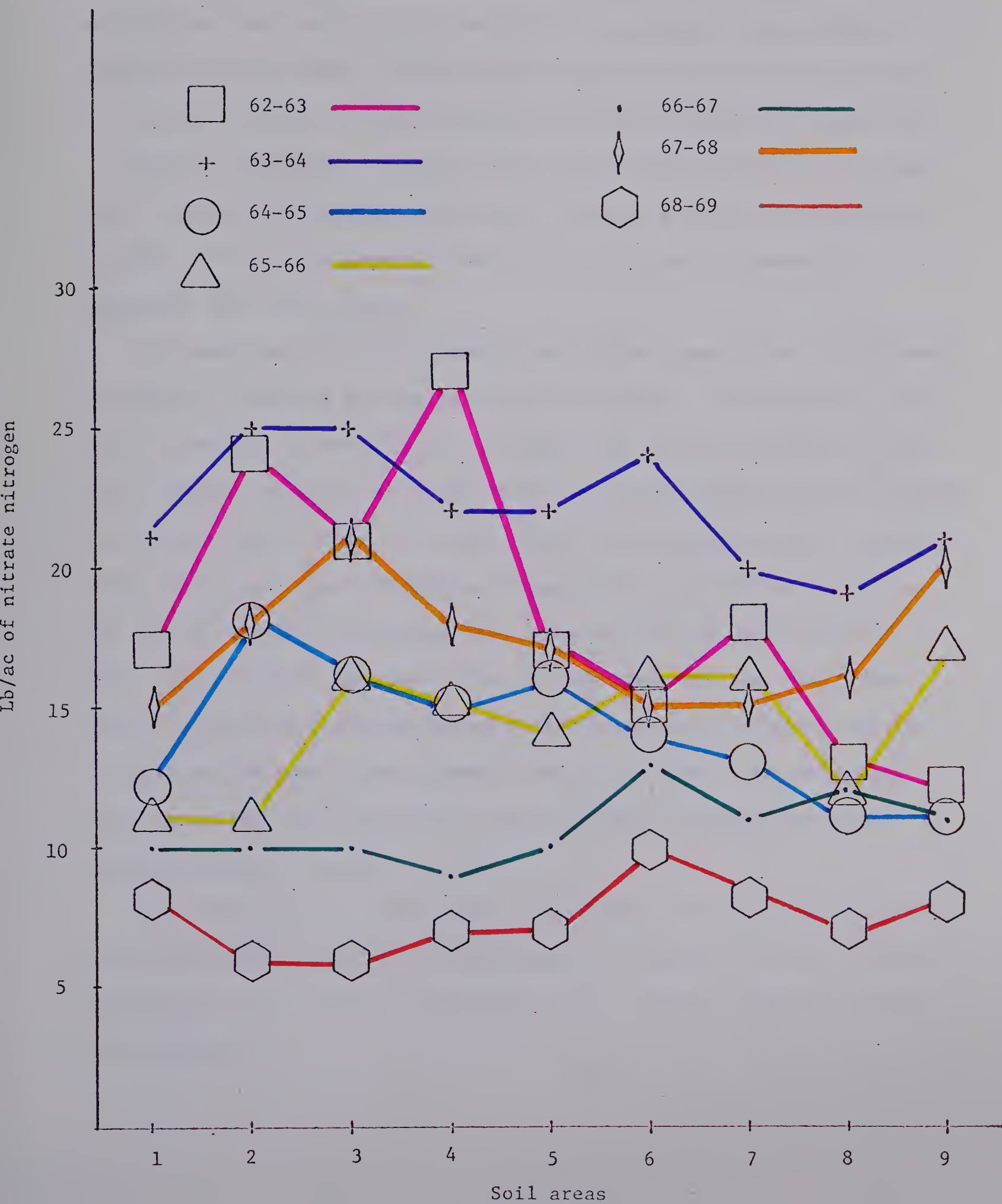








Figure 11. Nitrate nitrogen averages on stubble by sampling year and soil area. Months: Sept.-Apr.





nitrogen occurred. There is some interaction between the areas and years. As the graph points out, not all the areas respond similarly from year to year. If this were the case, the lines on the graph connecting the same sampling years would all be parallel or approximately parallel to each other. The literature emphasizes that nitrate nitrogen is probably the most variable nutrient with time and this is largely due to weather conditions. This fact alone will probably explain any of the major responses within an area by year. Despite a few noticeable interactions, however, the general trend is for all areas to respond similarly from year to year.

A comparison of both the stubble and fallow graphs shows a difference in relative levels of nitrate on fallow and stubble. For example, in the fall of 1966 the nitrate levels on stubble were lower than those for the fall of 1964. However, on fallow fields in the same areas just the reverse occurs, the 1966 nitrate levels are higher than the 1964 levels. In other words, there is a small difference between fallow and stubble levels in the fall of 1964 and a comparatively larger difference in the fall of 1966. Thus, contrary to the results already presented for Area 6 from Figure 9, fallow and stubble fields in all areas do not both necessarily increase and decrease simultaneously from year to year. As the graphs point out, stubble levels do not vary as greatly as fallow levels of nitrate from year to year.

(b) Phosphorus - - The effect of different crop types on the average value and standard deviation of phosphorus for Areas 2, 6, and 9 are shown in Figures 12, 13, and 14, respectively. The results, summarized briefly are as follows.



Figure 12. Comparison of phosphorus averages and standard deviations on different crop types for Area 2.  
Time period: All months, 1962-68.

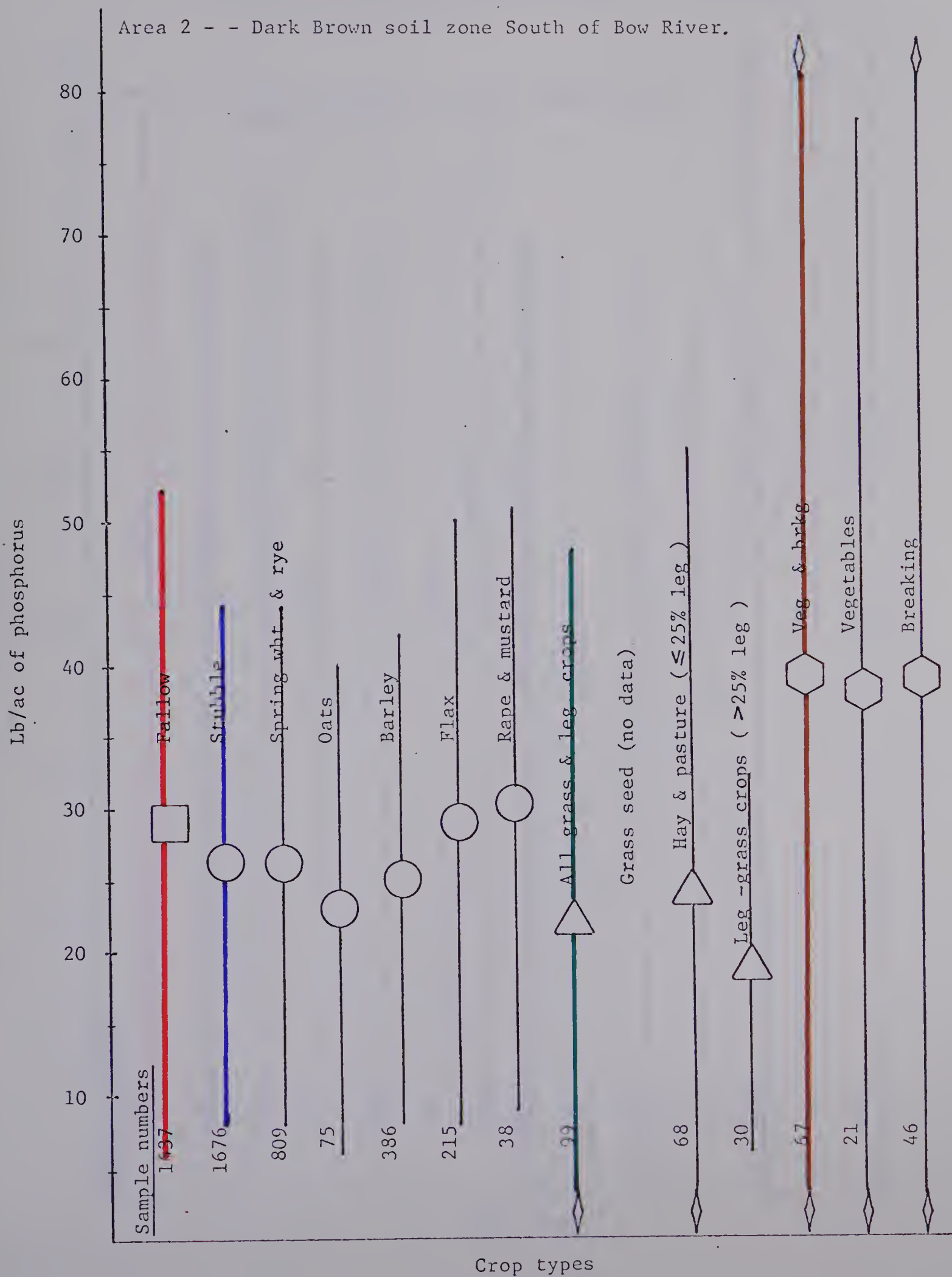




Figure 13. Comparison of phosphorus averages and standard deviations on different crop types for Area 6. Time period: All months, 1962-68.

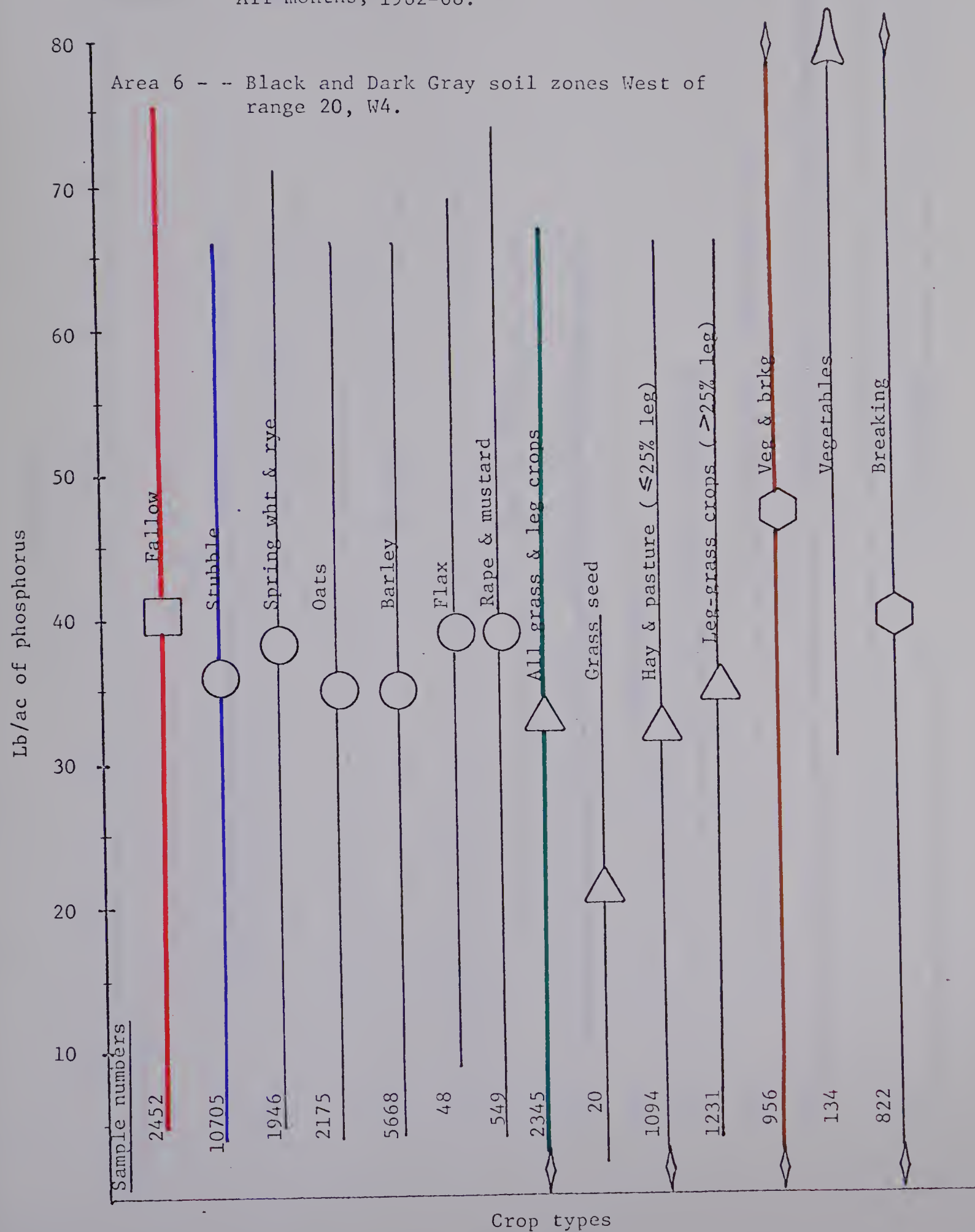
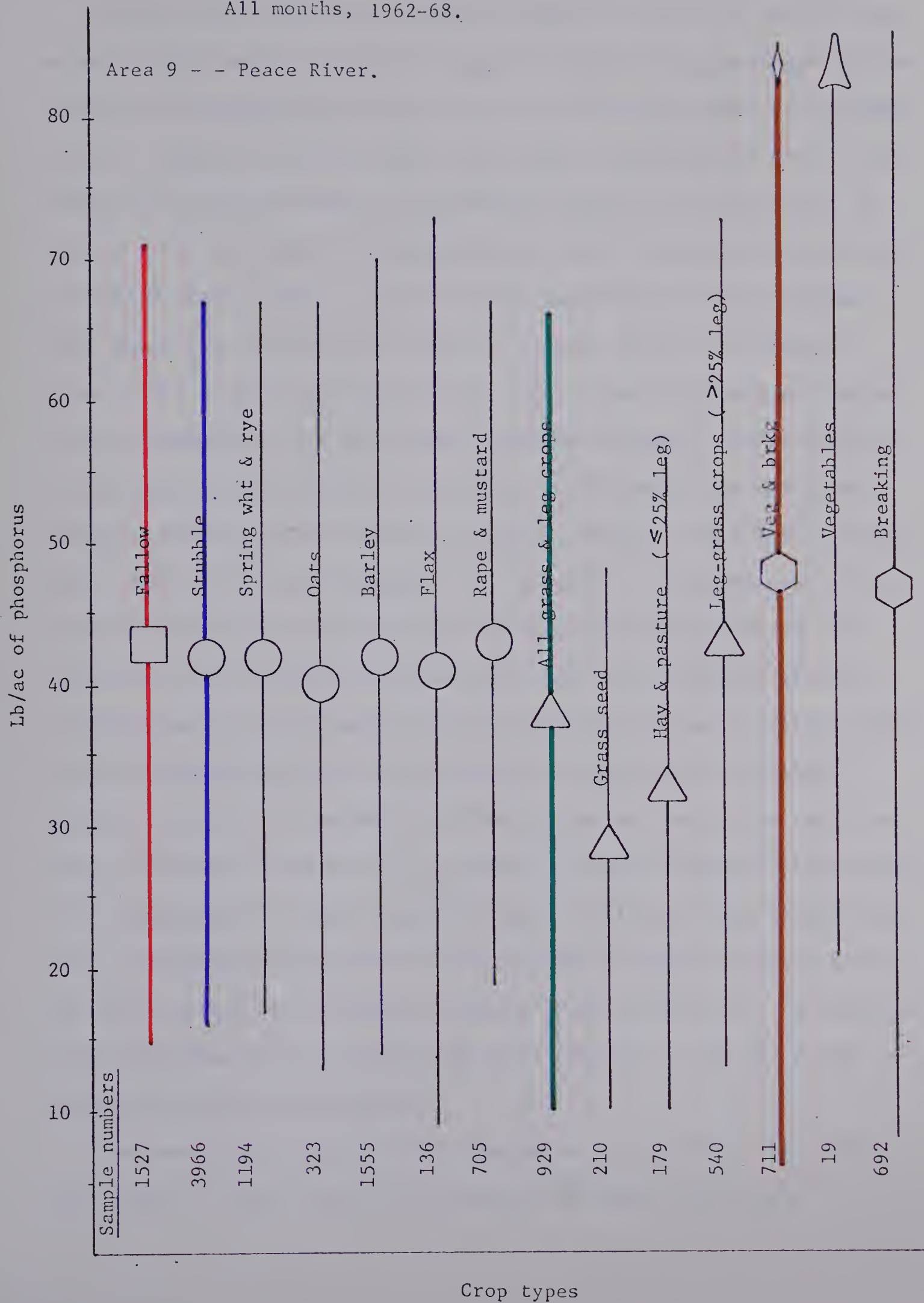








Figure 14. Comparison of phosphorus averages and standard deviations on different crop types for Area 9. Time period: All months, 1962-68.





Fallow fields tend to be slightly higher in phosphorus content than stubble fields which in turn are slightly higher than grass-legume fields. Fields with straight grass crops such as grass seed are very low in phosphorus. Land with breaking tends to be higher than fallowed land. Land cropped by vegetables shows an exceedingly high level of phosphorus in the top 6 inches of soil. In both Areas 6 and 9, the mean could not even be plotted on the graph. A more detailed inspection of the individual crops shows that land on which oats was cropped tends to be slightly lower on the average than other grain crops. Land previously cropped by rape and mustard, on the other hand, tends to be higher. However, with the prevalent, large standard deviations, no difference can really be implied. Generally for purposes of mapping, samples from fallow, stubble, grass-legume fields and breaking can be grouped. The almost unbelievable means and deviations on land cropped by vegetables would tend to indicate that these samples would be better left out. The high standard deviation usually accompanying the mean phosphorus value on breaking might also indicate its exclusion for purposes of mapping for two reasons. Firstly, inclusion of samples from breaking and not vegetables would require an internal adjustment in the present computer program, and secondly, new breaking tends to take place in fringe areas where other samples may not be available to compensate for any sporadic values present, and where the "edge" effect of the mapping function is most prevalent. In addition, the loss because of the exclusion of these samples is only 5% of the available samples on the average.

The results for the levels of phosphorus on different crops follow very closely to those reported by Tisdale and Nelson (op. cit.),



Riehm (op. cit.), and Parker et al. (op. cit.).

A comparison between phosphorus levels and standard deviations in different soil-climatic areas on fallow, stubble, and legume-grass crops is illustrated in Figure 15. The graph points out the previously noted slight decrease in phosphorus content as one goes from fallow to stubble to legume-grass fields. The standard deviations are of similar magnitude for fallow and stubble crops with those for legume-grass crops tending to be larger in all areas.

A study of the graph shows that as one goes from southern to northern Alberta there is a steady increase in phosphorus content in the top 6 inches of soil. The lowest area is Area 4 (21 lb/ac on stubble) in the southwest of Alberta, whereas the high area is Area 8 (49 lb/ac on stubble) in the northwest where Gray Wooded soils predominate. A comparison of the deviations from area to area on similar crops indicates a fairly close correlation between the mean values of phosphorus and the absolute deviation resulting in approximately equal coefficients of variation from area to area. Area 7 and 8, with large amounts of Gray Wooded soils, tend to be the most variable.

A comparison of the yearly variation of phosphorus on stubble and fallow in Area 6 is illustrated graphically in Figure 16. The results plotted in this manner point out approximately 3 to 5 lb/ac differences between fallow and stubble in the mean phosphorus content for Area 6. The deviations are similar, but with phosphorus values tending to be slightly more variable on fallow. The results show that levels of phosphorus fluctuate a noticeable amount from year to year. The peak year for Area 6 is the fall of 1965 with a stubble mean of 44 lb/ac. The lowest year is





Figure 15. Phosphorus averages and standard deviations by crop grouping and soil area. Time period: All months, 1962-68.

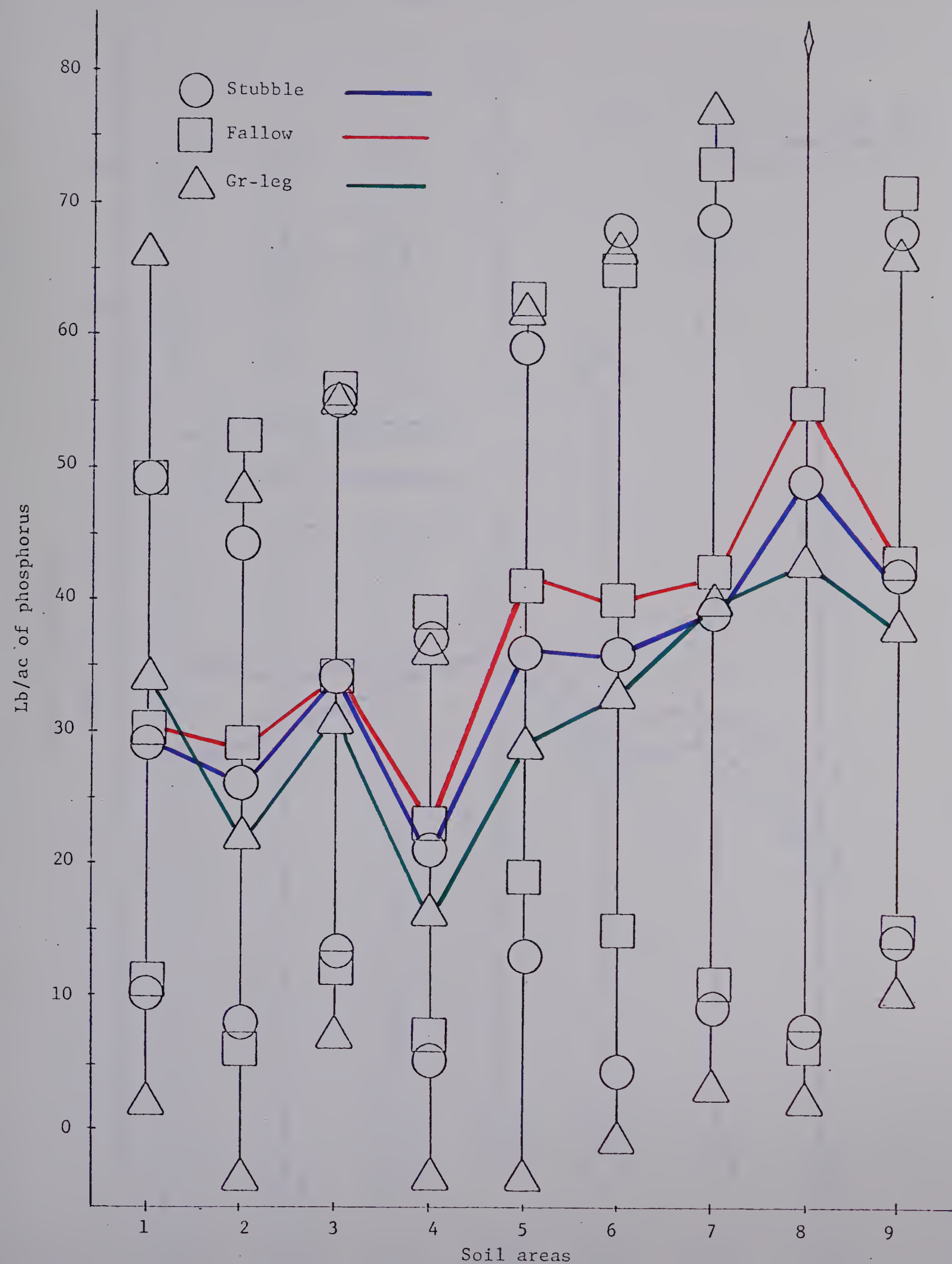
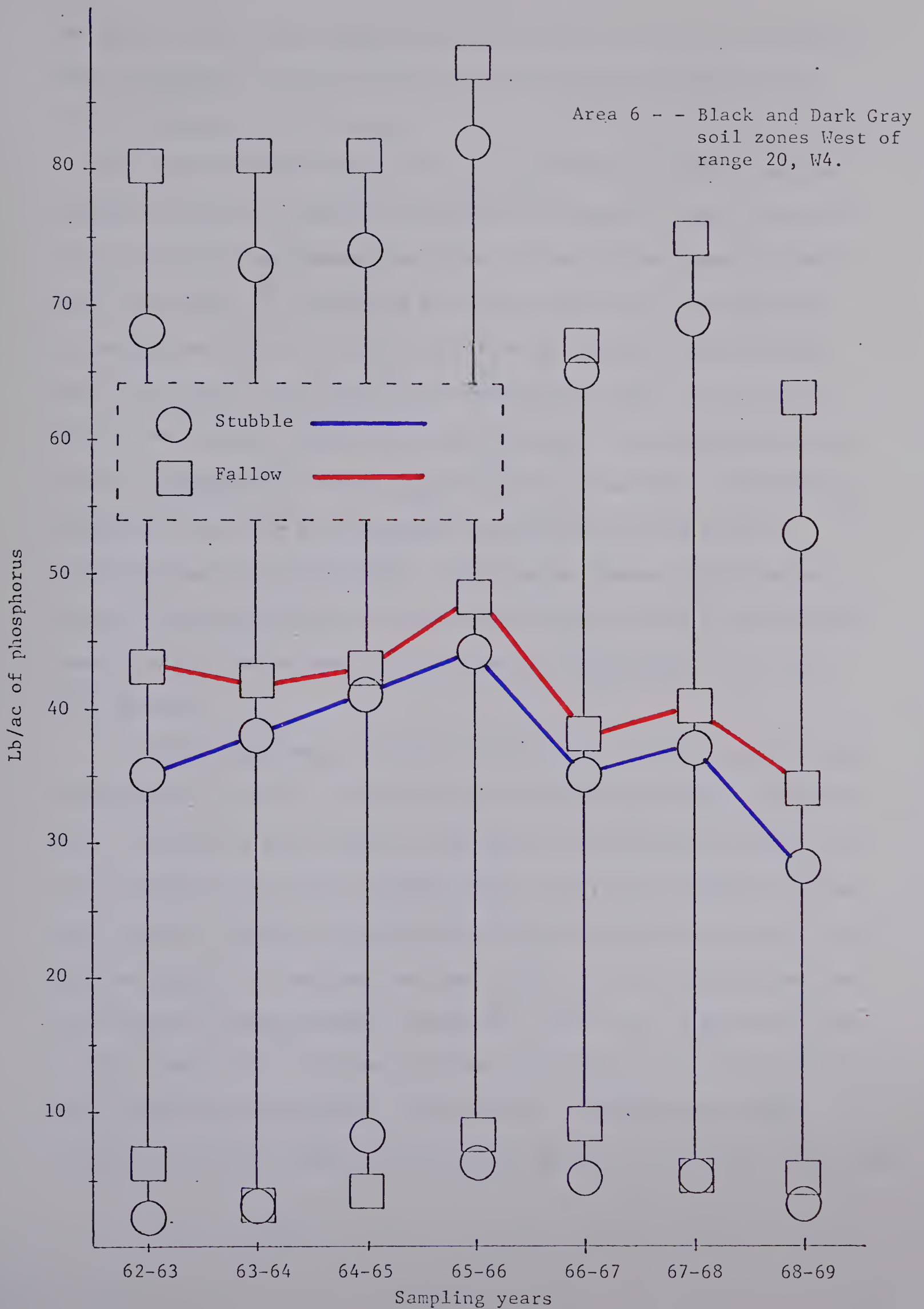






Figure 16. Yearly variation of phosphorus on stubble and fallow in Area 6. Months: Sept.-Apr. 111.





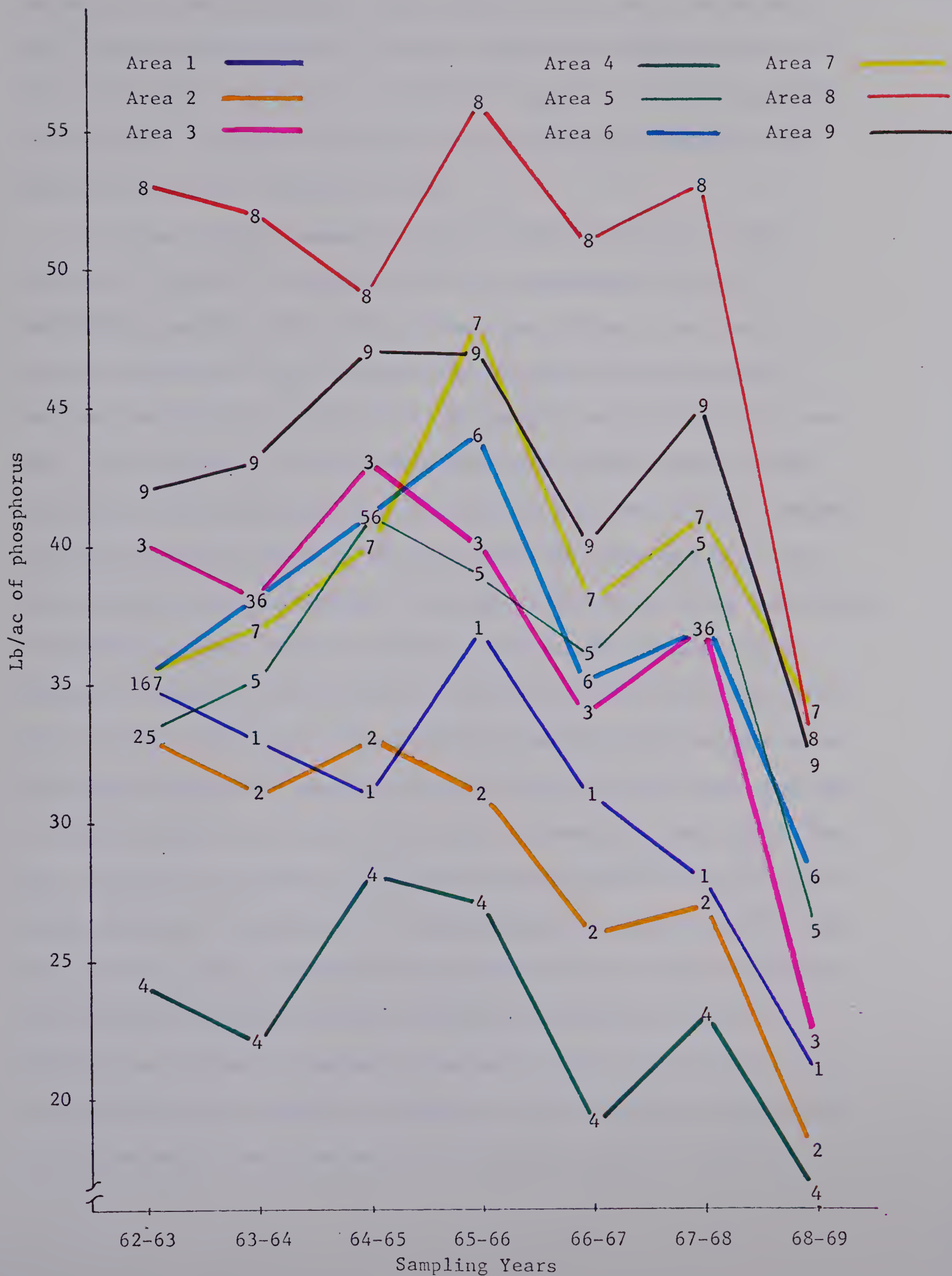
the fall of 1968 with a stubble mean of 28 lb/ac. There are no apparent large differences in the coefficients of variation for phosphorus in Area 6 from one year to the next.

The variation of the mean level of phosphorus on stubble from year to year for each soil area is illustrated in Figure 17. Here, one sees even more extreme variations than exist in Area 6. For example, Area 8 drops from a high of 56 lb/ac in the fall of 1965 to a low of 33 lb/ac in the fall of 1968 -- a drop of 23 lb/ac in 4 years. All the areas tend to vary from year to year, some more so than others. This is contrary to the results reported by Riehm (op. cit.) who states that after repeated sampling on the same fields year after year that no significant fluctuations occur in pH, phosphate, and potassium levels even on cultivated and fertilized fields. His results, however, might not be entirely applicable due to the fact that he measured only lactate-soluble phosphoric acid and because his study was only conducted over a period of 27 months.

A proposed suggestion for this variation from year to year has been the sampling. That is, from one year to the next an entirely different part or parts of a soil-climatic area might be sampled the heaviest. It is the opinion of the author, however, that this is very unlikely for two main reasons. Firstly, most districts within an area tend to send in proportional amounts of samples from year to year and although high and low fluctuations in sample numbers are bound to occur over an area from year to year, they usually occur on an almost random basis. An area is usually never sampled proportionately, but rather the more productive parts of the area are usually sampled the heaviest, and this is true year after year.



Figure 17. Phosphorus averages on stubble by soil area and 113. sampling year. Months: Sept.-Apr.







Secondly, if biased sampling is the cause for this yearly variation, then from any sampling year to the next some areas should increase and others decrease. However, as indicated in Figure 17, this is generally not the case. The graph shows that most areas respond similarly (but at different rates) from year to year.

One other possible cause for yearly variation might be yearly Laboratory variation. Vermeulen (1953), an agronomist from the Netherlands, points out that many factors can influence the results in routine analyses and that a mathematical treatment should accompany analyses periodically to check their reliability and comparability over time. The relocation of the Alberta Soil Testing Laboratory from the Agriculture Building (University of Alberta) to the new Dr. O.S. Longman Building (University Farm) in the late summer of 1968 might be a contributing factor in the low soil phosphorus and nitrate values encountered in the fall of 1968. Table XV with the data for the province shows a decrease on stubble from 39 lb/ac in the fall of 1967 to 27 lb/ac in the fall of 1968. According to the available literature this appears to be an abnormal decrease in phosphorus content from one year to the next and it would be easy to say that a part of it is laboratory variation. However, despite the available literature (of which there is very little on yearly phosphorus variation) it is believed by the author that the main reason for any yearly variation in phosphorus content is due largely to the difference in annual weather conditions affecting the rates of fixation and release of available phosphorus in the top 6 inches of soil. In addition, recent laboratory analyses on stored samples has shown very little laboratory variation as a result of a change of location.





The results in Figure 17 should be examined critically. The graph emphasizes that the levels of phosphorus are more influenced by soil areas than by yearly variation. In addition, with the large standard deviations that exist for phosphorus, it is difficult to assess the degree of importance of the yearly variation. Phosphorus content is more closely associated with soil types than with changing weather conditions. Nitrate levels, on the other hand, are more closely related to prevailing weather conditions.

(c) Potassium - - Figures 18, 19, and 20 present a graphical comparison of potassium means and standard deviations on different crop types for Areas 2, 6, and 9, respectively. The graphs show that there is generally very little difference in levels of potassium on fallow and stubble crops. This is contrary to the summary for the province in Table XV and further emphasizes the possible slant that can be brought about by crop type and sample number relationships. Farrari and Vermeulen (op. cit.) state that the average potassium content of grassland is higher than that of arable land in the Netherlands. Comparable results can only be found in Area 6. Levels of potassium on breaking are similar to those on fallow and stubble crops. The levels, in general, follow a similar crop-effect pattern as found with phosphorus except for breaking and vegetable averages. The coefficients of variation for the potassium levels on all crops are much lower than those for either phosphorus or nitrate nitrogen. Breaking tends to have the largest standard deviation. For purposes of mapping all crops can be grouped.

A comparison of potassium levels and standard deviations by soil area on fallow, stubble, and grass-legume crops is presented in Figure 21.



Figure 18. Comparison of potassium averages and standard deviations on different crop types for Area 2. Time Period: All months, 1962-68.

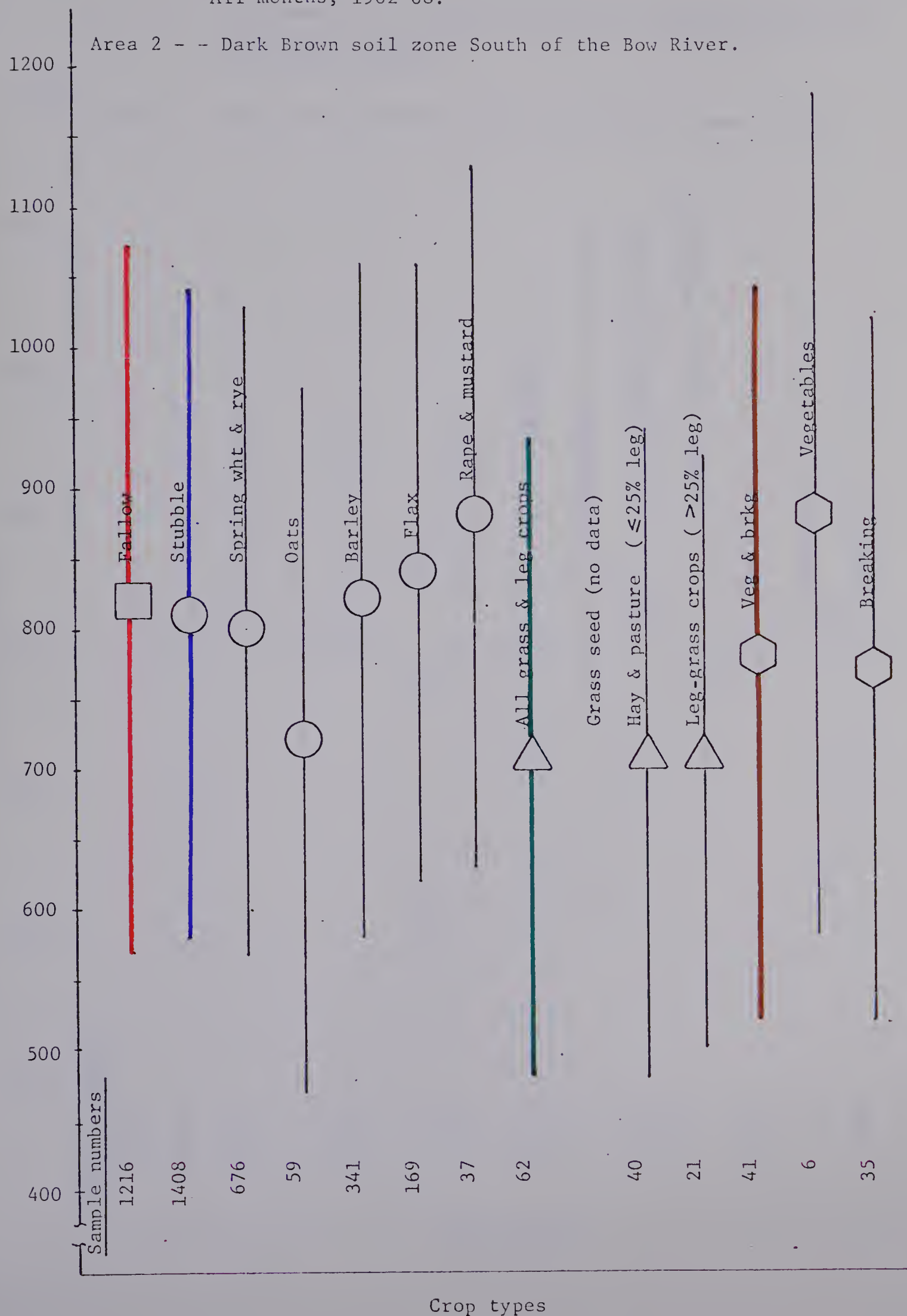




Figure 19. Comparison of potassium averages and standard deviations on different crop types for Area 6. Time period: All months, 1962-68.

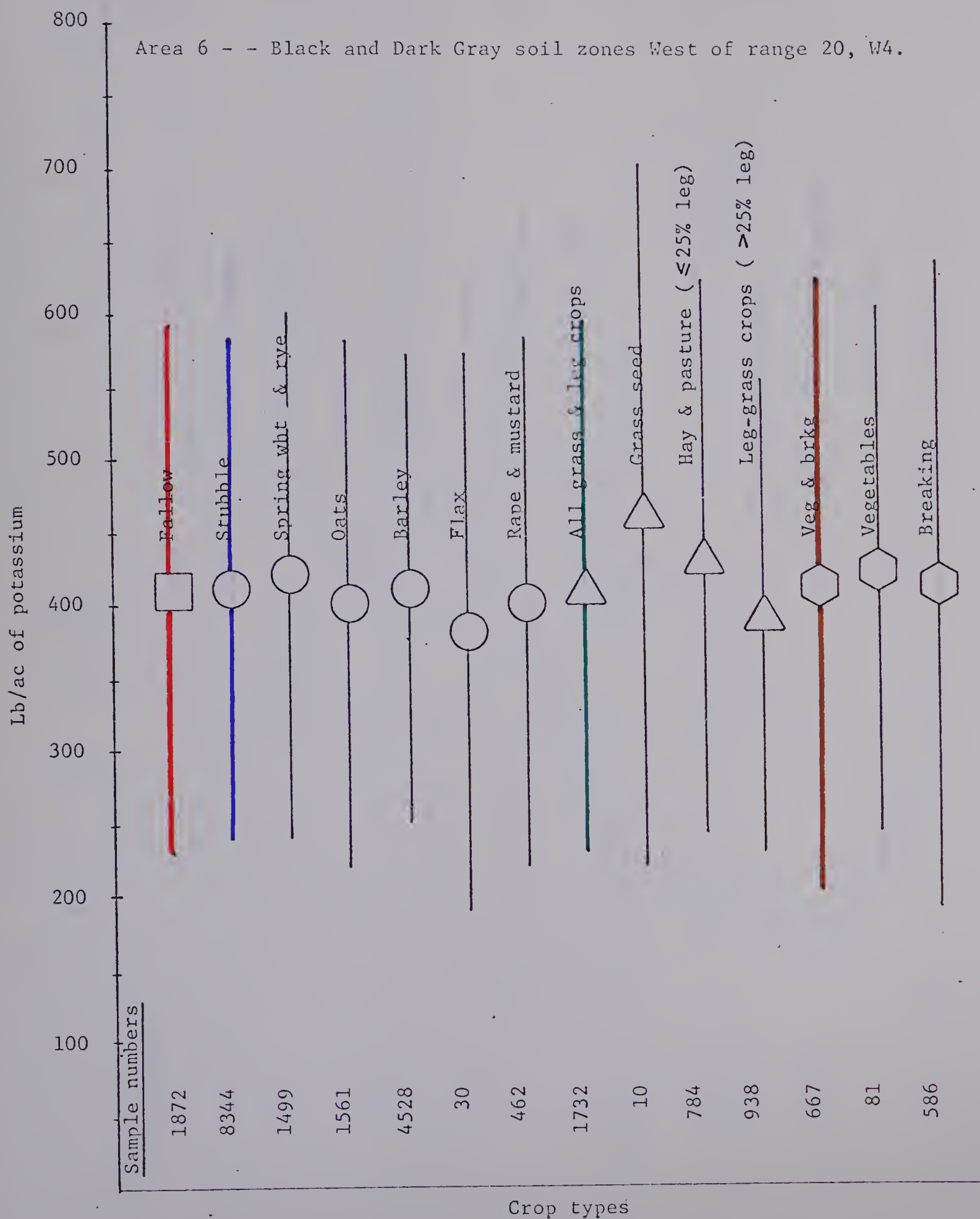




Figure 20. Comparison of potassium averages and standard deviations on different crop types for Area 9. Time period: All months, 1962-68.

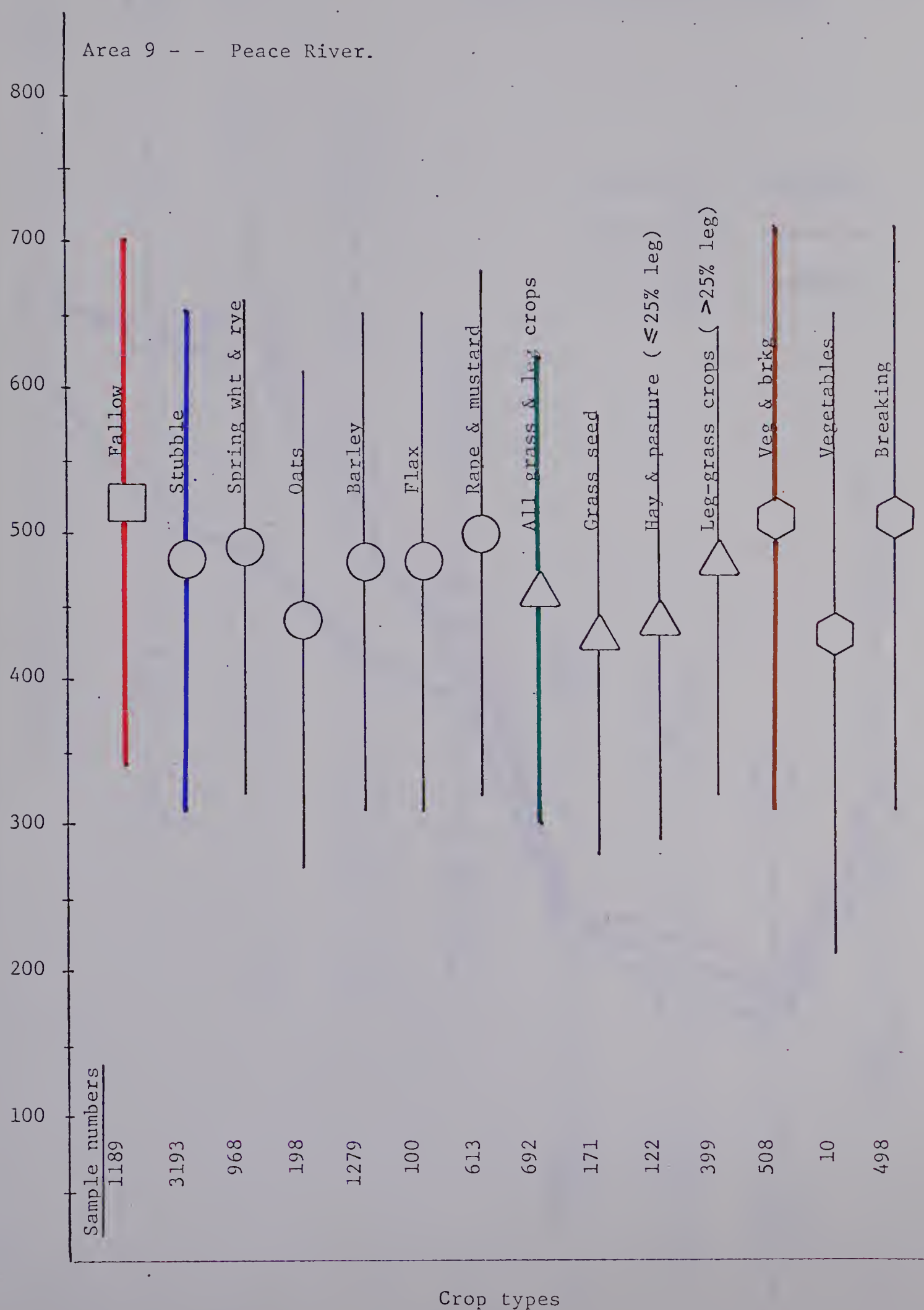
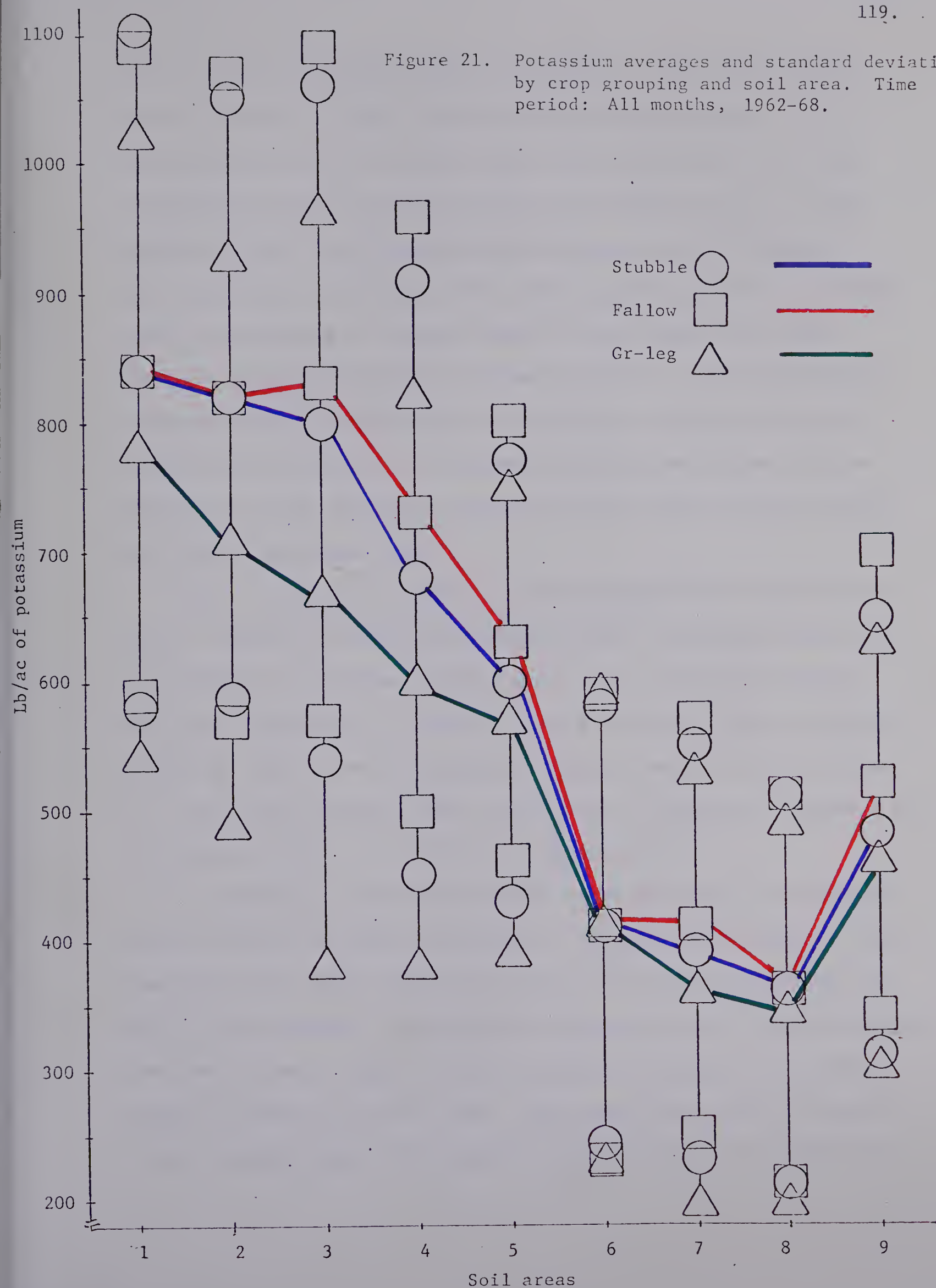






Figure 21. Potassium averages and standard deviations by crop grouping and soil area. Time period: All months, 1962-68.





Fallow tends to be slightly higher in potassium content than stubble which in turn has a higher content than grass-legume crops. This is particularly true in the southern areas where grass-legume crops are considerably lower in potassium content than stubble crops. An examination of Table XVI, illustrating the average crop percentages by area, shows that in the areas where a wide separation exists, the grass-legume crop grouping is composed mainly of grass crops rather than crops having high percentages of legume. However, a comparison of the potassium averages on grass crops versus legume crops does not give a definite trend this way. The standard deviations associated with the means of the three crop types indicate that very little relative variation between crop type exists.

There is a distinct decrease in the potassium levels from Area 1 to Area 8. Areas 1, 2, and 3 are relatively high in potassium content with over 800 lb/ac on stubble. Areas 4 and 5 show a decrease to 680 and 600 lb/ac, respectively. Areas 6, 7, and 8 indicate a sharp decrease to around 400 lb/ac. Area 9 then shows a climb to about 480 lb/ac. Areas 6, 7, and 8 have slightly larger coefficients of variation as compared to the southern areas, i.e. 0.4 and 0.3, respectively.

A comparison of yearly means and standard deviations for potassium levels in Area 6 on fallow and stubble is illustrated in Figure 22. The graph shows very little difference in the mean levels of potassium for the four years tested. No noticeable differences occur in the variations from year to year. Figure 23 shows the yearly averages of potassium on stubble for each of the soil areas. The southern areas tend to have the largest absolute change, but relative to their mean this is probably not



Figure 22. Yearly variation of potassium on stubble and fallow in Area 6. Months: Sept.-Apr.

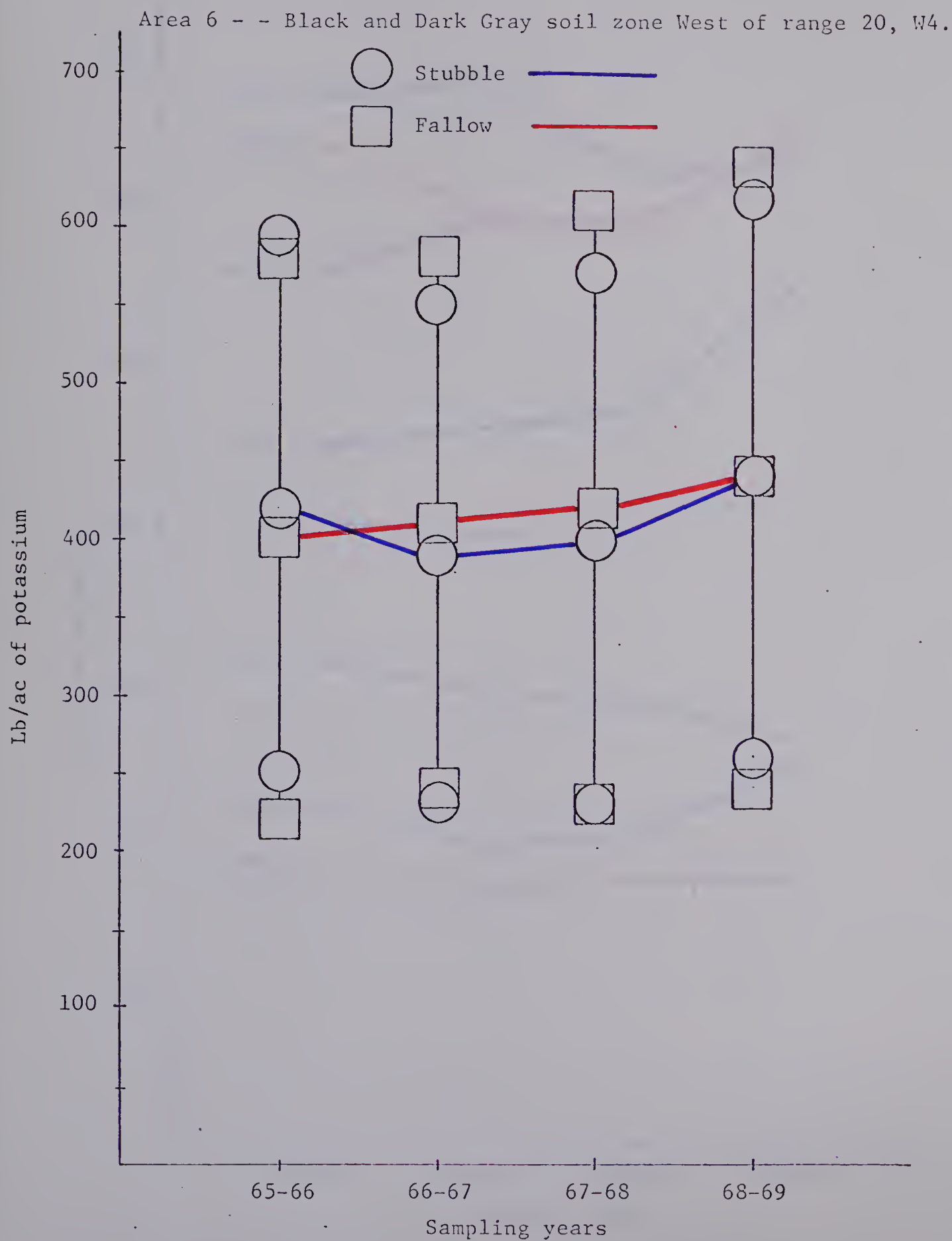
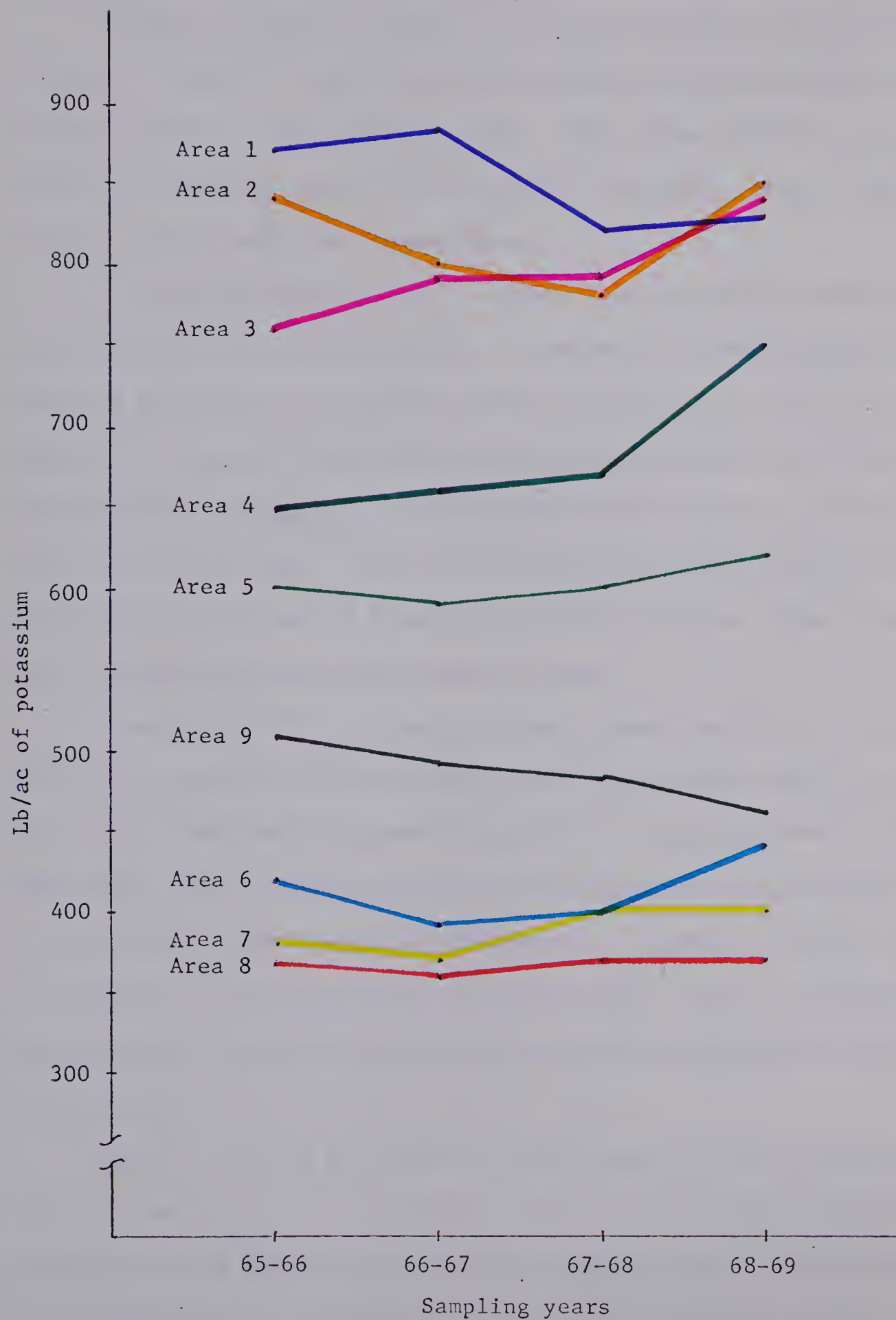




Figure 23. Potassium averages on stubble by soil area and sampling year. Months: Sept.-Apr.







very significant. Area 4 has the largest change, an increase from 650 lb/ac in the fall of 1965 to 750 lb/ac in the fall of 1968.

Generally, Figure 23 emphasizes two major features about potassium. Firstly, it does not change considerably with time, and secondly, it is closely related to soil-climatic area. This agrees with the results quoted by Riehm (op. cit.), Parker et al. (op. cit.), Nelson (op. cit.), Goettel (1962), and other researchers.

(d) Soil reaction - - pH - - Of the four laboratory determinations studied, pH is the least variable. A comparison of pH averages and standard deviations on different crops for Areas 2, 6, and 9 is shown in Figure 24. The graph shows that no significant differences usually exist between major crop groups. The majority of the means are within  $\pm .2$  pH units of each other. Very little difference can be detected in the magnitude of the standard deviations between different crops. The variation on breaking tends to be slightly larger.

A comparison of the pH averages and standard deviations in the different soil areas on fallow, stubble, and grass-legume crops is found in Figure 25. The graphical presentation in this figure serves to illustrate again the equality of the pH means under different crop groupings. Areas 2 and 5 tend to show the major differences. Generally, stubble crops are about .1 pH units higher than fallow crops. This is a reversal in the trend noted by most other researchers who have found pH to be lower on cropped land.

Areas 1, 2, and 4 in southern Alberta tend to have the more basic pH's with means around 7.1 pH units. Areas 3, 5, 7, and 9 display the lowest pH's with means of approximately 6.4 pH units. The coefficients of variation do not change noticeably from one area to the next.



Figure 24. Comparison of pH averages and standard deviations on different crop types for Areas 2, 6, and 9.  
Time periods: All months, 1962-68.

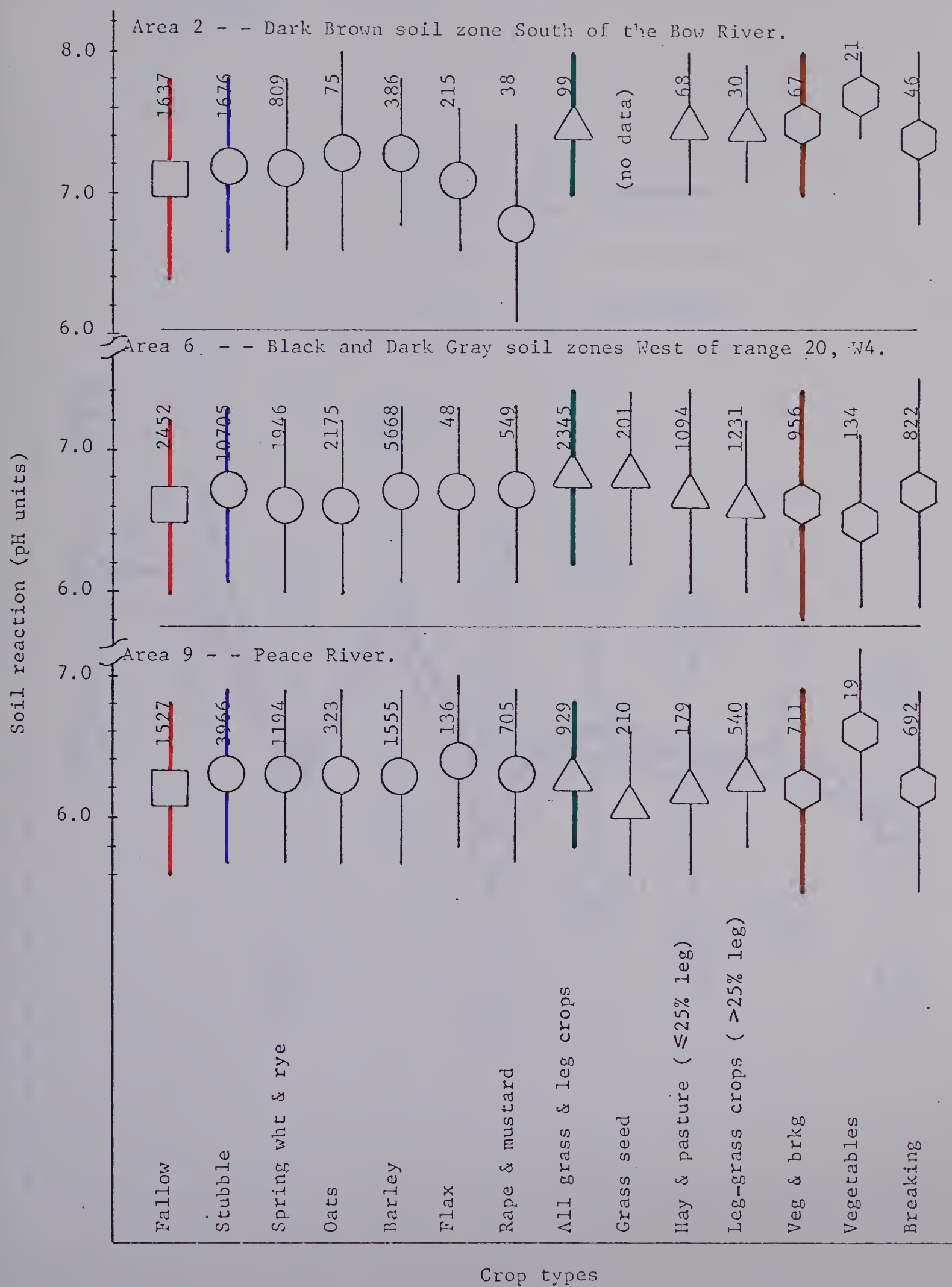




Figure 25. Comparison of pH averages and standard deviations by crop grouping and soil area. Time period: All months, 1962-1968.

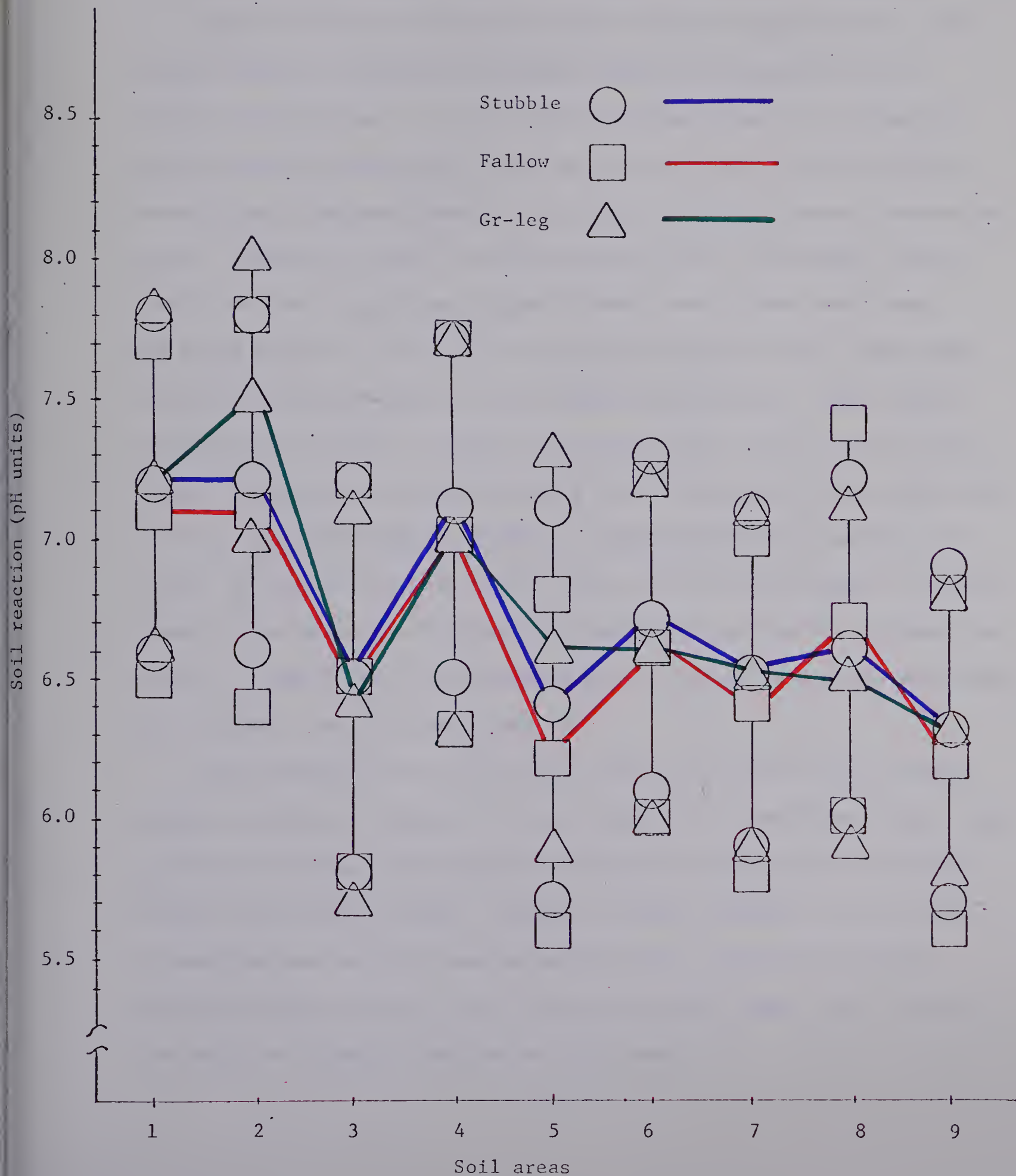






Figure 26 shows a graphical representation of pH means and standard deviations over the years for stubble and fallow crops in Area 6. The graph shows very little annual change in the means and standard deviations.

Figure 27 shows the pH means by soil area and sampling year. The results show the previously noted gap between the pH means for the southern areas (Areas 1, 2, and 4) and the other areas. The change in the mean pH value from year to year is most noticeable in the southern areas, where a maximum change of .4 pH units can occur between consecutive years. Generally, however, considering the scale of the graph, there is usually no real significant changes between years in most soil areas. The graph points out that pH is more closely related to soil areas than to yearly conditions that may cause slight fluctuations. Small annual fluctuations do occur, of course, but usually they affect all areas similarly. For example, the graph shows a drop of pH in all areas in the fall of 1967. The following year there is a corresponding rise in pH for all areas. It should be noted, however, that not all areas change at similar rates or even direction in every year; although the trend is certainly to do that. Areas 1 and 2, for example, seem to change in opposite directions for each year from 1962 until 1965.

Most researchers have found small annual variations in pH; however changes as much as 1 pH unit have been noted. Bell and Thornton (op. cit.) and Raupach (1951a, 1951b) both state that these variations are not consistent from place to place. Contrary to their findings, the pH results by sampling year and soil area generally show a consistent variation (direction-wise) from one year to the next for all areas. Soil reaction, obviously, is closely linked to the soil areas.





Figure 26. Yearly variation of pH on fallow and stubble in Area 6.  
Months: Sept.-Apr.

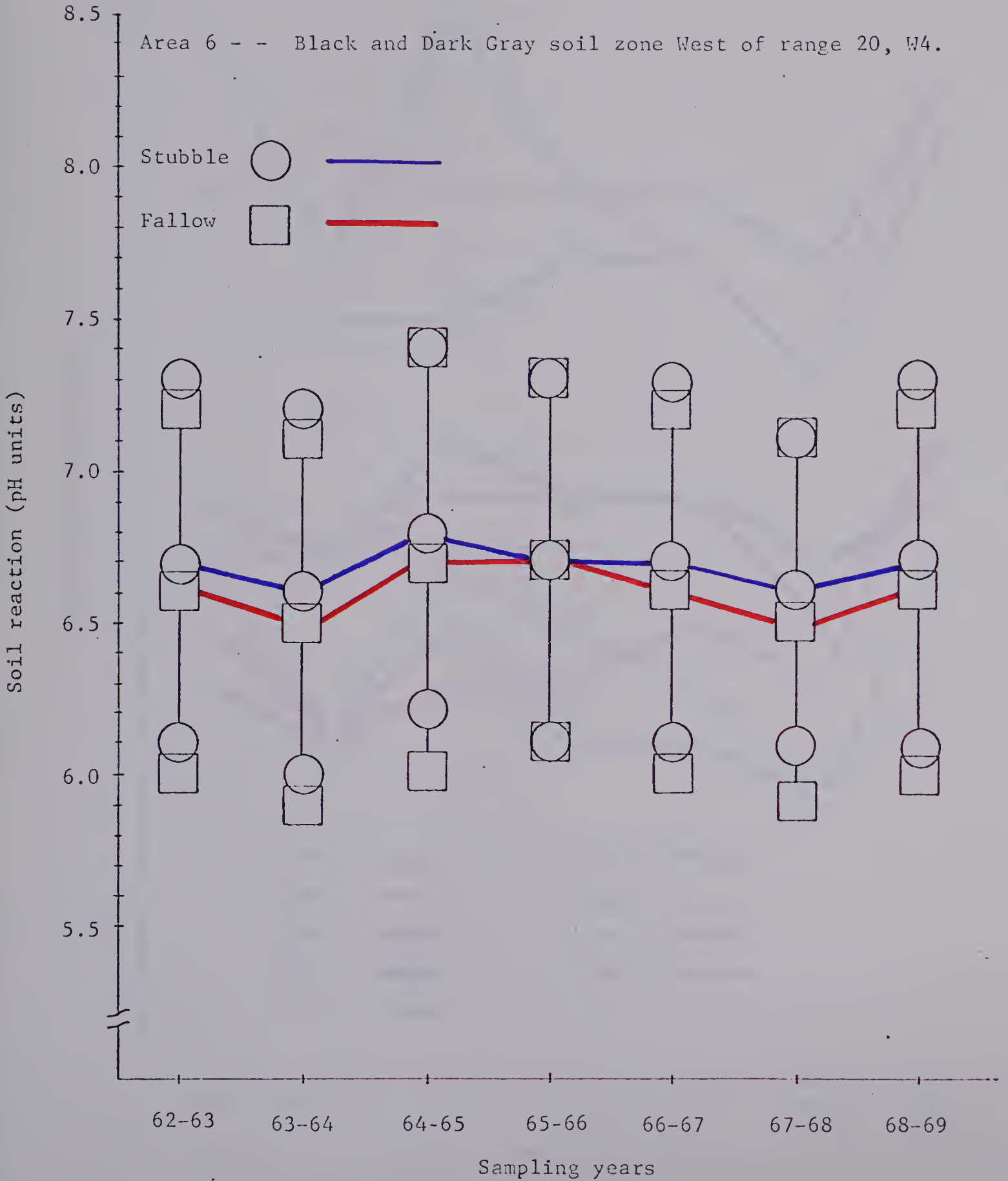
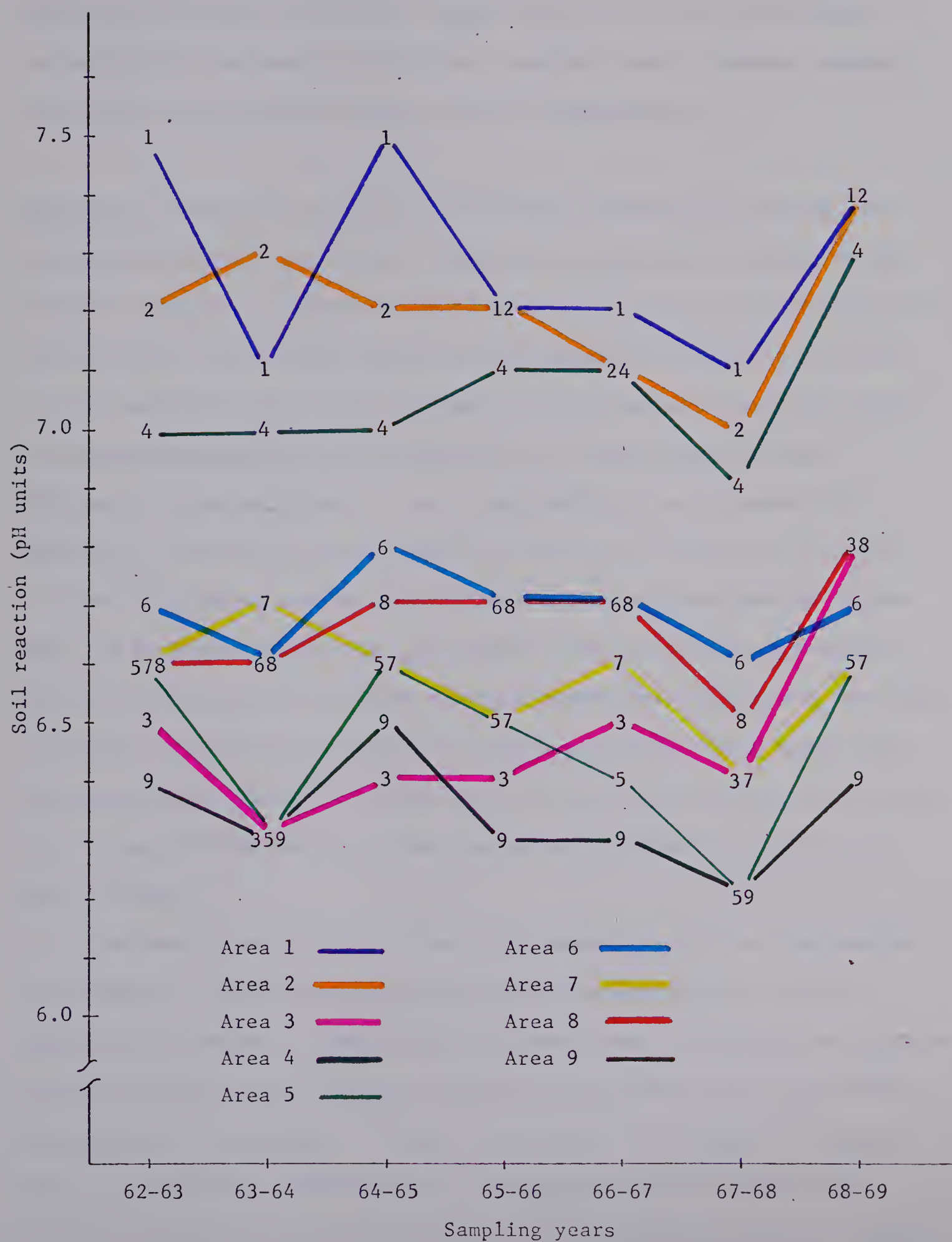




Figure 27. Averages of pH on stubble by soil area and sampling year. Months: Sept.-Apr.





The results of the area summaries for pH point out that pH can be mapped on all crops and over all years because very little difference exists between the means on different crops and years. However, mapping the results for each year might prove very interesting.

Counties: Summaries for each of the Counties, Municipal Districts, Improvement Districts, and Special Areas were conducted as outlined in the section "Methods". A numerical tabulation and presentation of these summaries would be exceedingly long and time consuming, making it difficult to group general trends. The easiest to follow and most practical method of summarizing such data is by displaying the mean values for each district on a map of Alberta. Such a map was previously presented in Figure 1 b. The main disadvantages of this type of presentation is that no idea of variation can be shown without additional complication to the map; and as Parker et al. (op. cit.) point out, soil types and associations do not necessarily follow district boundaries. Therefore, this type of summary is sometimes difficult to relate to specific soil types. The advantage is, of course, a simple, easy-to-follow presentation of the data, where areas of high and low values can easily be located, at least on a general scale.

A necessary prerequisite to any such undertaking, is an examination of the number of samples representing the mean for each district, thus, giving an indication of the accuracy of such a mean. Although the districts are not the same shape, they are sometimes of a similar size to the critical circle (37 townships). Thus, an estimation of the number of samples that are required for 80% accuracy can be gotten from Table XIII. As Figure 28 points out, the majority of the cultivated areas are well sampled,



but the fringe areas are usually very poorly sampled.

Figure 28 is a map of Alberta showing the number of soil samples received by district for all months and crops from 1962 to 1969. Any samples with greater than 20 acres represented by one core were eliminated in this study (3%). The resulting map, then, gives a fairly good idea of the sampling density throughout the province. The areas surrounding Edmonton, Red Deer, Calgary and Lethbridge are the most densely sampled areas. Municipal District No. 90, directly north of Edmonton, is the most densely sampled district with over 2500 samples received in the above time period. The Peace River area is fairly well sampled in the arable parts. The non-arable areas and fringe areas have the lowest number of samples. The area north of Medicine Hat and east of Drumheller appears to be the most sparsely sampled of the cultivated area.

The code "X" has been used in all district maps to designate less than 10 or no samples existing, with the latter being more predominant. Where less than 10 samples were present, irregular values tended to exist, and therefore a requirement of 10 samples was arbitrarily set before a value was recorded.

The area summaries, discussed previously, show large differences between nitrogen on fallow and stubble. Figures 29 and 30 show the county and district averages of nitrate nitrogen on fallow and stubble, respectively. The averages are summed over all the months for the years 1962 to 1969 in both maps. A comparison of the two maps points out the previously noted difference between fallow and stubble in all districts. Nitrogen on fallow is fairly high (41-50 lb/ac) in the Edmonton area, but tends to drop to lower values (21-30 lb/ac) in the southern areas. The Peace River



















region has similar nitrate values on fallow as does the Edmonton area.

The nitrate levels on stubble follow a similar trend to those on fallow, but do not vary as much. In fact most stubble nitrate levels rarely exceed 20 lb/ac, with the majority of the counties and districts having between 11 to 20 lb/ac of nitrate.

A map of Alberta illustrating the phosphorus averages per district is shown in Figure 31. Soils of the southern part of the province are inclined to have low values of phosphorus (21-30 lb/ac). A northward movement across Alberta shows a more or less steady increase up to the districts northeast of Edmonton where a fairly sharp jump occurs and high values (51-70 lb/ac) occur. The Peace River region also contains some fairly high values of phosphorus (41-50 lb/ac). The area to the southeast of Edmonton appears to have an almost random occurrence of coded 3 and 4's, making any trends within that area unclear.

Potassium averages on stubble for the districts is shown in Figure 32. The map obtained follows closely those trends already explained for each of the soil areas. There is a steady and fairly steep decrease in potassium content as one goes from southern Alberta to northern Alberta with the areas to the northeast of Edmonton containing the lowest values (201-300 lb/ac).

Averages for pH on stubble for each district with 10 or more samples are shown in Figure 33. The soils of southern Alberta are the most basic soils with pH's around 7.1 to 7.5. The Calgary-Red Deer area has pH's of 6.6 to 7.0. The Edmonton area seems to be dominantly 6.1 to 6.5 in pH with similar averages existing for most of the Peace River districts. To the northeast of Edmonton the pH rises slightly, about 6.6 to 7.0.







Figure 31. Map of Alberta showing phosphorus averages on stubble in counties and districts. Time period: All months, 1962-69.

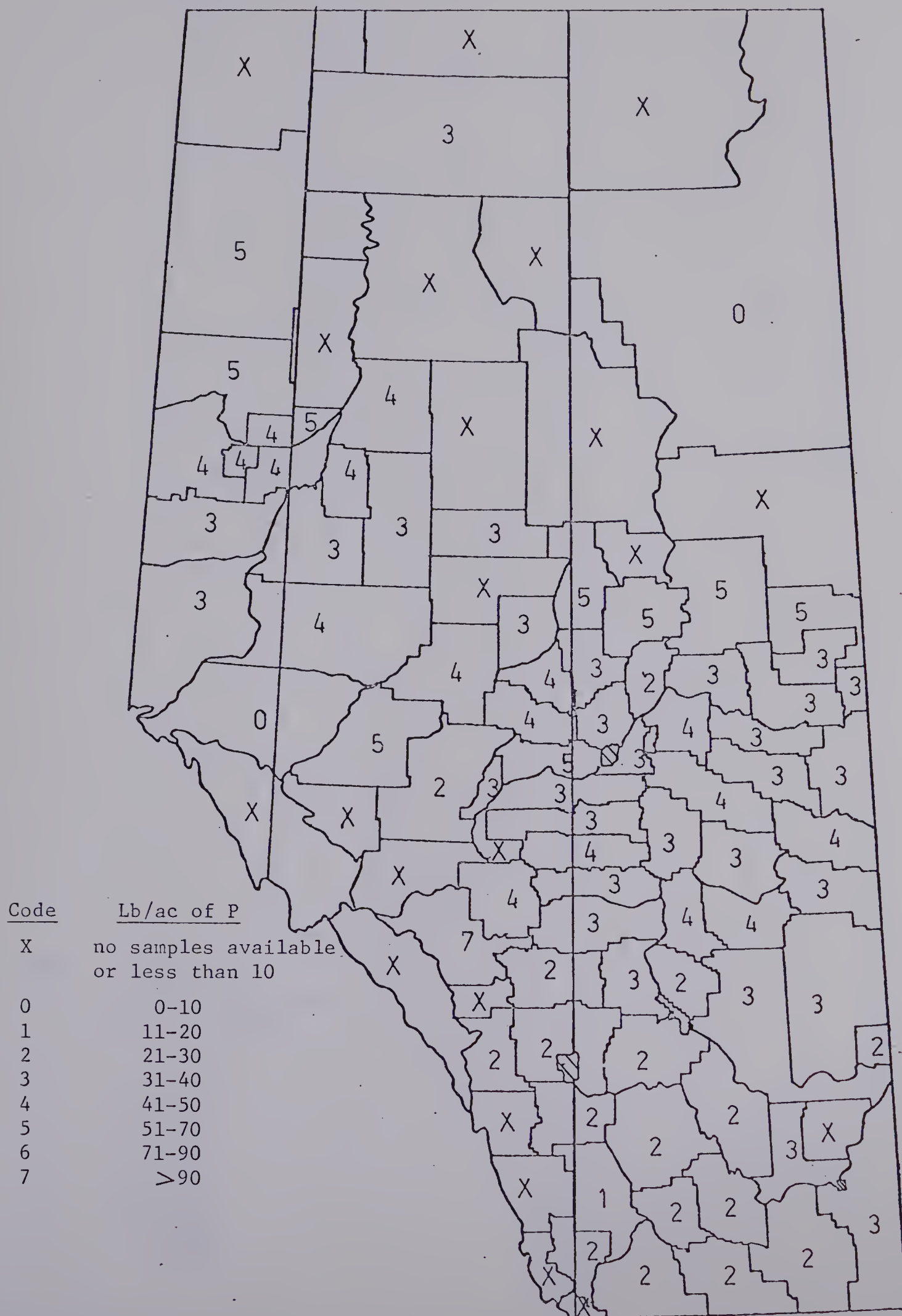
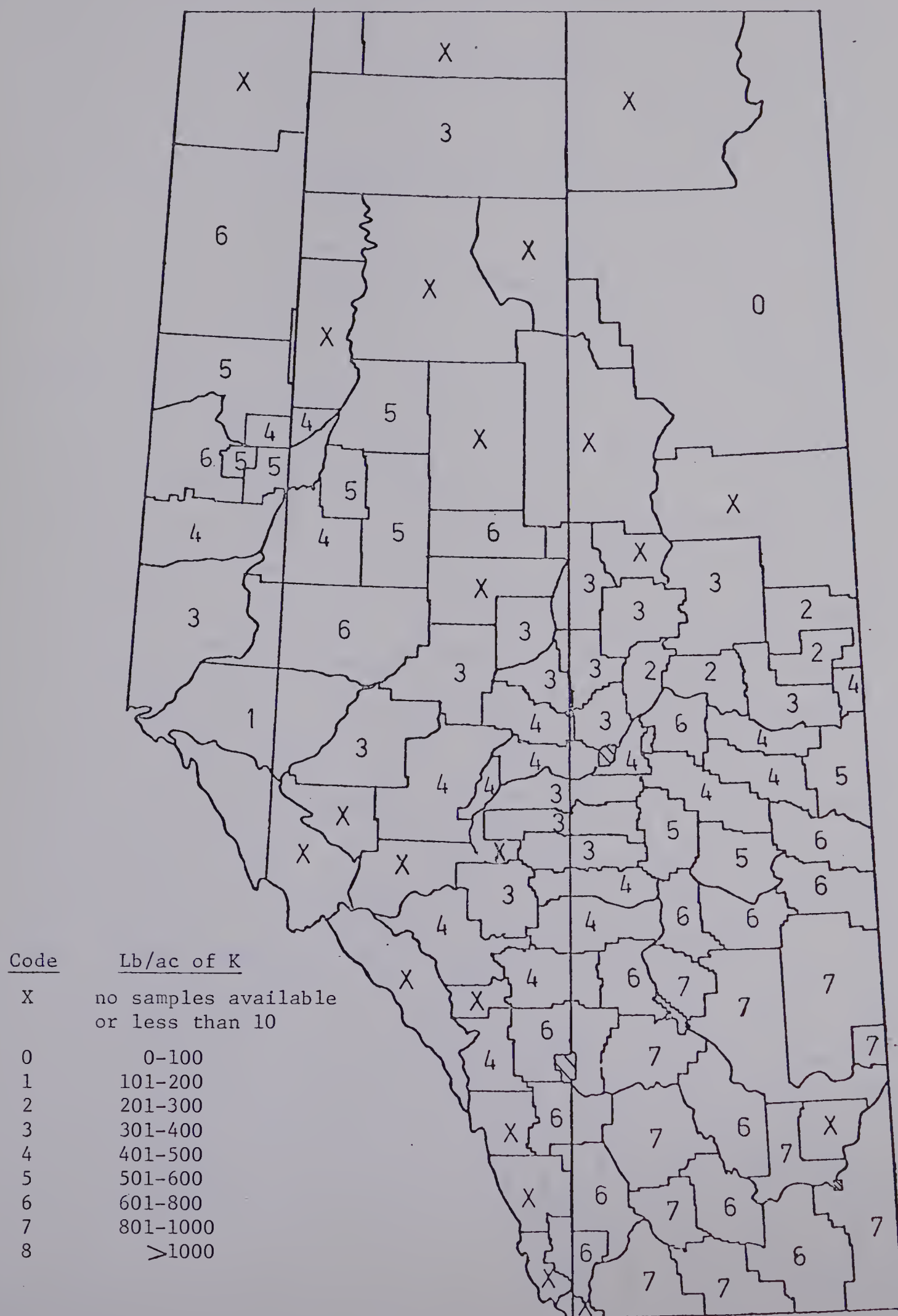




Figure 32. Map of Alberta showing potassium averages on stubble in counties and districts. Time period: All months, 1962-69.









The results tend to follow closely those previously presented in the area summaries for all four nutrients. Despite some disadvantages, the district summaries presented in this form give a concise and uncomplicated picture of the general nutrient status of Alberta soils in which definite trends prevail.

### Computer Mapping

Results of the methods used: Before the actual mapping technique selected to map the soil test data for the province was chosen, several varied and different techniques were tried and the results and ease of application to the A.S.F.T.L. data were recorded. A brief summary of the results for each of the techniques will be given below. Both the area and data selected for this study, as well as a short discussion of the actual mapping technique may be found in the section "Materials and Methods".

(a) "Eyeballing" - - Figure 34 shows the area in which the trial mapping runs were made with the values of phosphorus rounded-off into appropriate intervals and displayed for each township. The "O"'s mark the location of 59 townships which have no samples. At first glance one can note the confusion that exists in trying to map such data where extremes exist in adjacent townships. Many "hills" and "hollows" exist as a result of local variation.

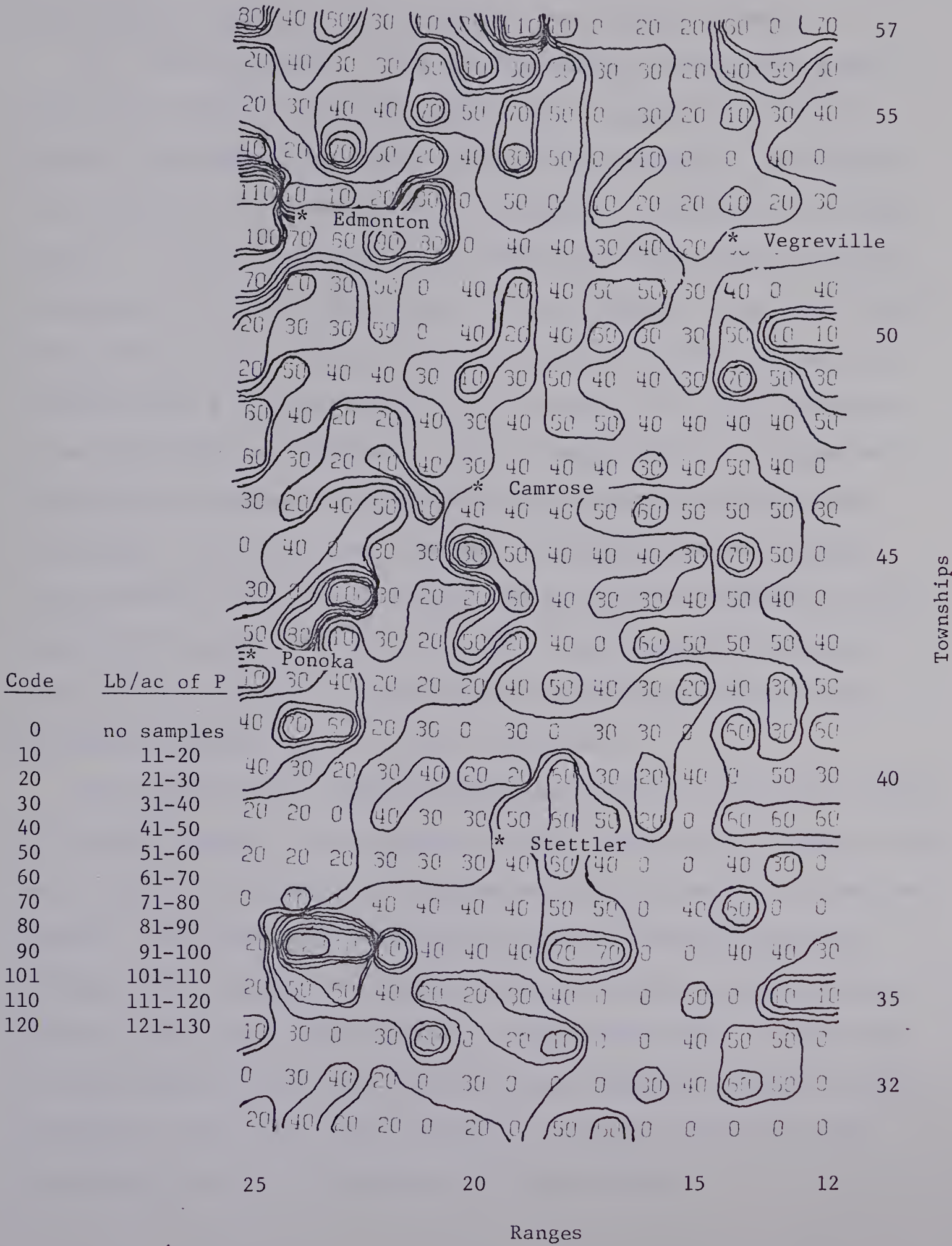
With the aid of the electronic computer the data has been presented in a map form and the process of "eyeballing" in what appears to be the trends can begin. The method from this point on is no longer analytical but related to "know-how" and past experience with the data. The procedure







Figure 34. Map of trial area displaying original values of phosphorus rounded-off and contoured.





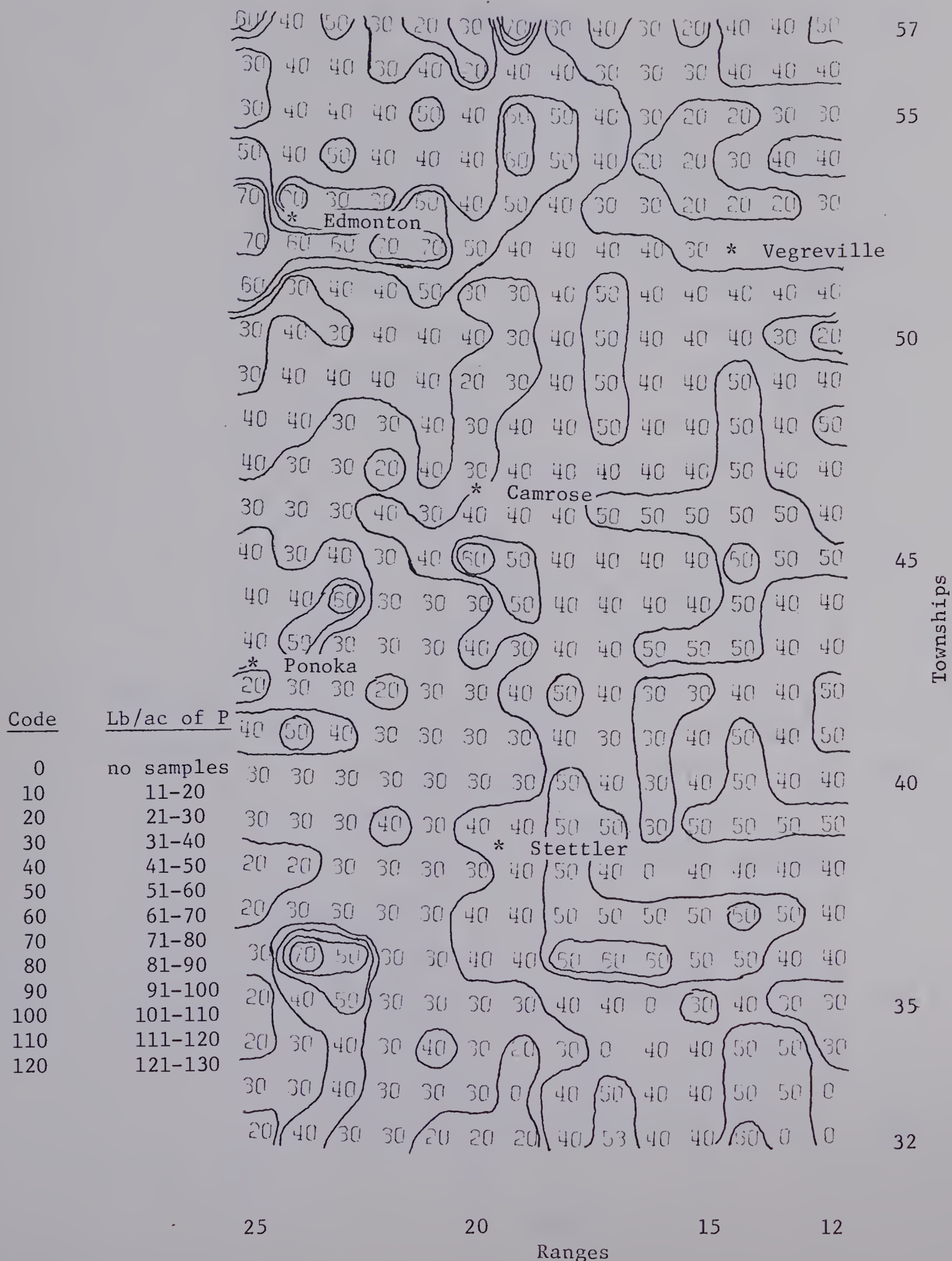
from here on would be to manually outline broad areas of approximately equal phosphorus content; thus presenting smooth region contours.

(b) Profile method - - The profile method and variations of the method were applied to the trial mapping area. Figure 35 is a map of the trial area showing one of the results of this technique. The length of the profile or the weighted moving average was 7 townships with weightings of 5, 3, 2 and 1 assigned to the central and three nearest townships, respectively, in all four directions. An exact diagram of the cross-shaped mapping function used is shown in Appendix IIa. At least 5 samples were required within a critical radius of one township in all four directions before a map value was calculated. Thus 5 samples had to be present in 5 townships (including the one for which the map value calculation was being made). The results show that only 6 townships were not assigned values because of the lack of samples. However, the 53 townships that previously had no samples were predicted for, illustrating the predicting power of this technique. The prediction is based on the values of the surrounding townships in the x- and y- directions.

The results of the weighted profile method show a considerable degree of smoothing compared to the original rounding in the previous figure. The many "hills" and "hollows" have disappeared and larger regional trends have appeared. The results are considerably easier to comprehend and more pleasing to the eye than the observed values. In many instances, however, the contours tend to have what is known as a "fingering" effect. Rather than a circular smooth contour, the contours have a tendency to assume a finger-like shape. Most likely, this attribute is a result of lack of a poly-directional focus as is obtained in the circle method.



Figure 35. Map of trial area displaying results of the profile method.







The technique is readily transformed into a computer program and easily applied to the A.S.F.T.L. data.

(c) Circle Method - - Figure 36 is a contoured map of phosphorus representing the application of the circle method to the trial mapping area. The map shows the results of the circle method where 4 concentric circles have been used with the weights of 4, 3, 2 and 1. The inner two concentric circles (the central township and 8 surrounding townships) acted as the critical circle and at least 5 samples were required within this critical circle before a prediction was made. Only one township did not meet this requirement.

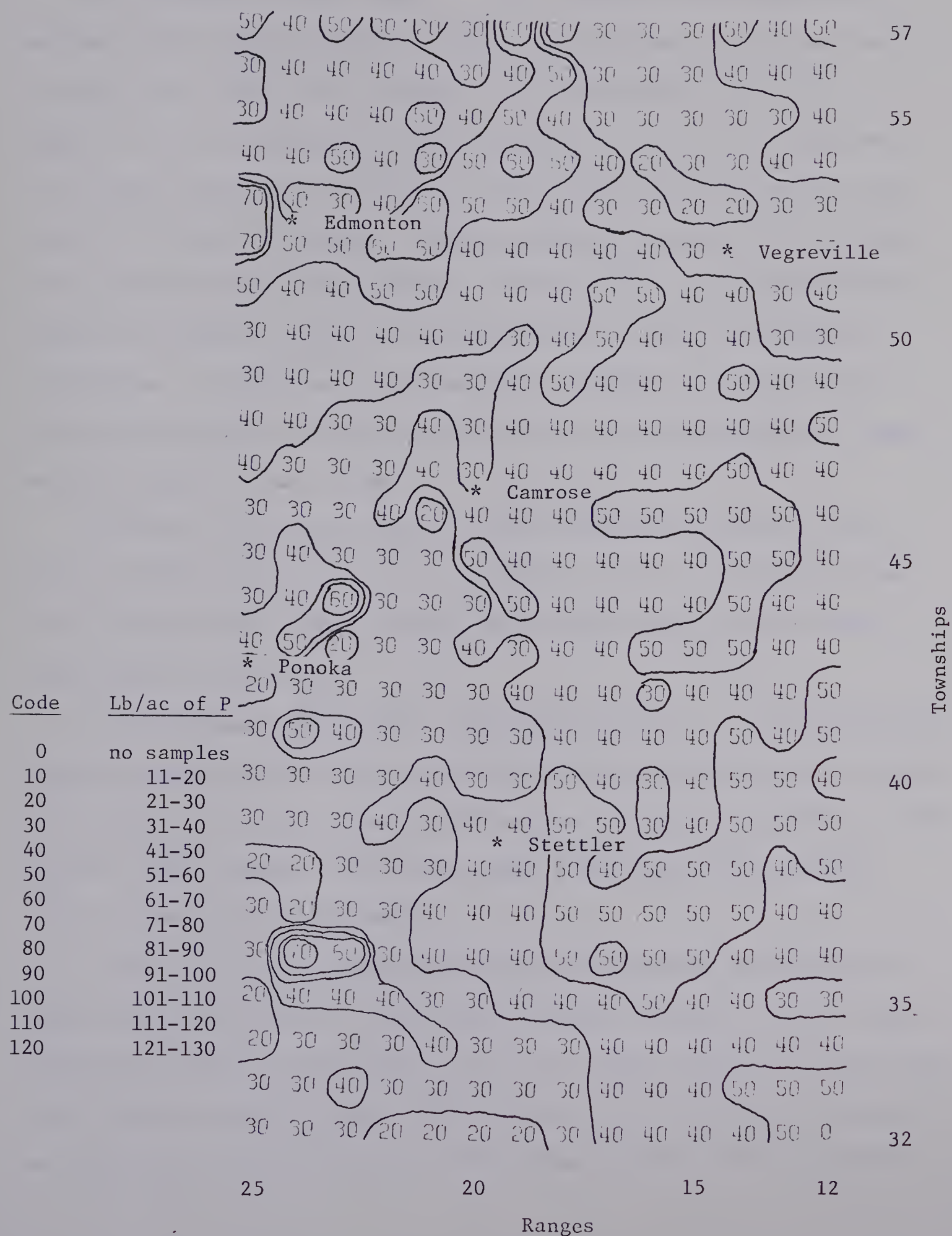
The method has eliminated the majority of the irregular minor fluctuations caused by what Krumbein (1956) considers local effects. The resulting regional picture is a relatively smooth surface with gradients in various directions. The local component in the map may be defined as the difference between the observed values and the smooth regional form. The local component by itself would display a spotty pattern of "hills" and "hollows".

Both the profile and the circle method give similar results, the circle method tending to give smoother results, producing a clearer regional picture. The two techniques are easily applied in a computer program and can be used efficiently and effectively to map the soil test data. Both techniques can be weighted and can have a controlled length or diameter and a critical distance or radius. However, the one fact that would favour the circle method over the profile method is that it is probably the more accurate averaging technique of the two, because it takes into account samples in all directions, not just in 4 directions. It is a more complete method in this respect, and this may well be critical in areas of low sampling density,





Figure 36. Map of trial area displaying results of the circle method.





especially with regard to preciseness of the final map.

(d) "Expected value" method - - The "expected value" method was applied to the trial area using averages for individual cell values and ignoring those cells with no samples. The final result is shown in Figure 37. The technique has indeed smoothed the data, but in doing so it has left a very spotty map which is quite different from the maps produced by the other techniques. Krumbein (1956) states that this technique should only be used on maps that are relatively free from complexities inasmuch as it takes no account of interaction terms that are sometimes significant. In addition, it automatically includes all direct polynomials in the regional surface, and hence it is not overly useful for large maps with many rows and columns.

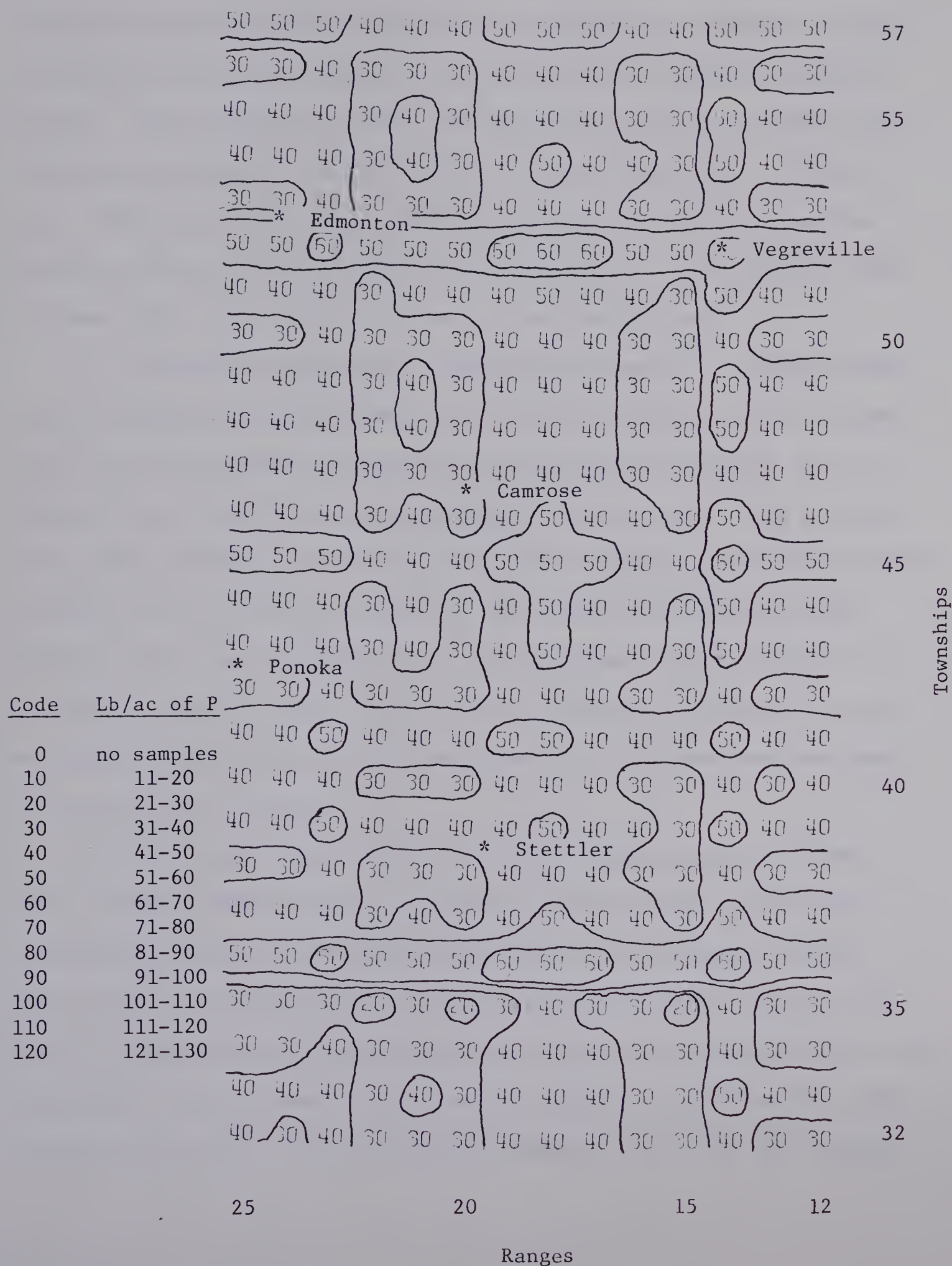
Thus, not only do the results of this technique eliminate it from further consideration for mapping the A.F.S.T.L. data, but definitely the size of the province (using townships as the basic mapping units) would cause grave inaccuracies.

(e) Orthogonal polynomials - abbreviated method - - The abbreviated regression method differs from the "expected value" method in that it allows the selection of the more important direct row and column polynomials. The "expected value" method, on the other hand, automatically includes all direct row and column polynomials.

Several trial maps were produced using this method, selecting different direct polynomials. The results, however, proved to be quite fruitless with no visual relationship to any of the other techniques except that of the "expected value" method. When all direct row and column polynomials were used, the final map showed an almost exact replica of the "expected



Figure 37. Map of the trial area displaying results of the "expected value" method.







value" method, as should be the case. Although the technique is stated to be suitable for fairly rapid analysis of maps with numerous rows and columns, the results do not show this. In addition the technique requires considerably more programming than any of the other techniques presented thus far. This technique, as well as the complete regression method using orthogonal polynomials, requires all data equally spaced and no missing data. Results show that it is probably more suitable for areas smaller than the province and is therefore quite unadaptable to the province wide soil test data, where townships form the basic mapping unit.

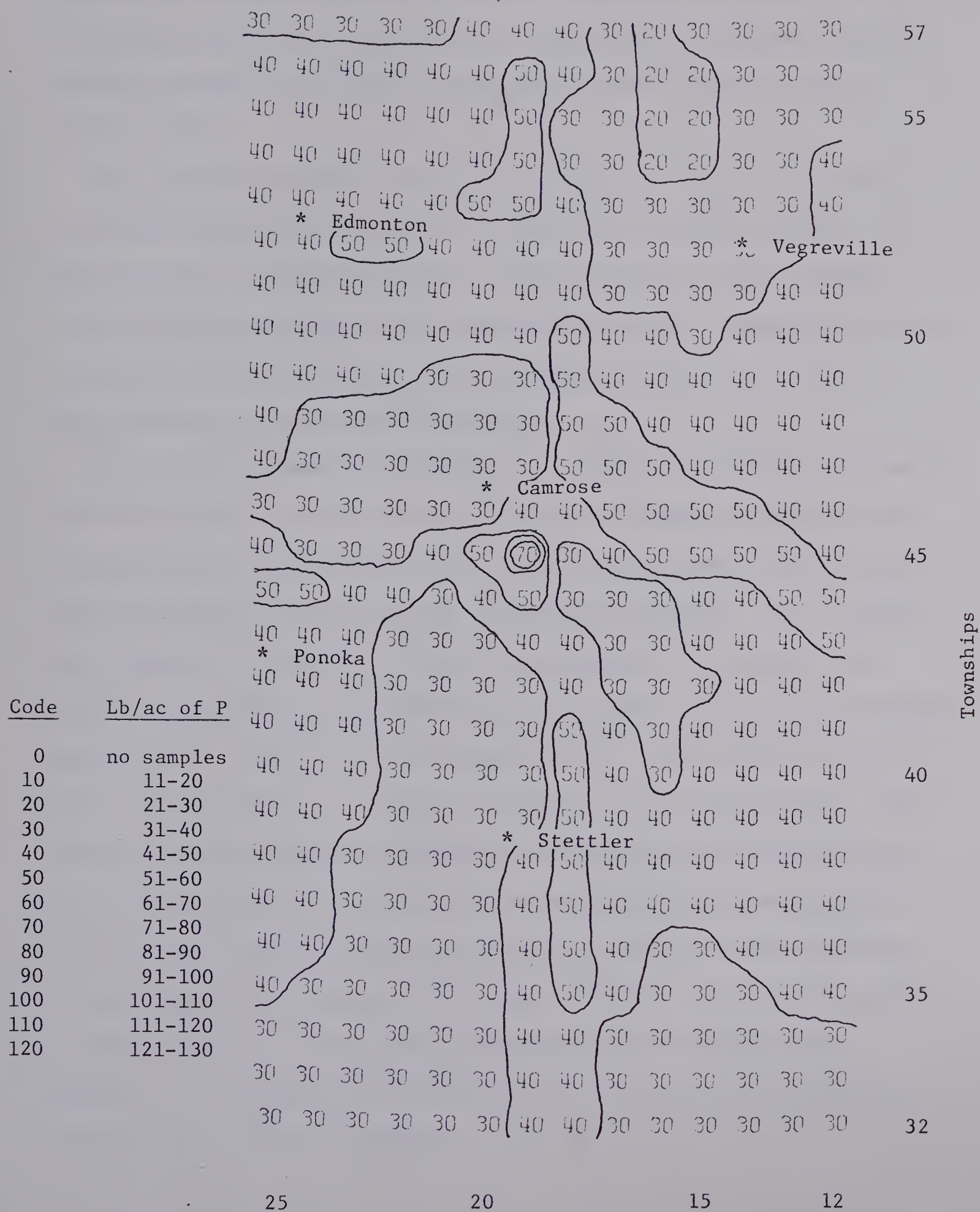
(f) Orthogonal polynomials - complete regression - - The one advantage the complete regression method has over the other methods is the fact that it not only allows the selection of all direct polynomials, but in addition any of the interaction coefficients. One result of the application of this technique to the data of the trial mapping area is illustrated in Figure 38. This map is a result of the complete regression method, whereby, direct row polynomials up to the 6th power, direct column polynomials up to the 4th power, and almost all interaction polynomials up to the 5th power for the row and the 3rd power for the column have been used to derive the final result.

Of all the maps illustrated for the trial mapping area, this map yields the most general picture of the phosphorus content in the area. A comparison with the circle method shows that it has somewhat similar trends, but is far too general and lacks what may be called "accuracy of detail". The inclusion of more direct and interaction polynomials into the mathematical function used to describe the regional response surface would bring out more detail. Most researchers, however, have found that usually





Figure 38. Map of the trial area displaying the results of the complete regression method using orthogonal polynomials.





no more than the cubic terms are necessary to adequately describe the regional trends, this being especially true with geophysical data. Contrary to the majority of the results, it appears that a much more complex function is needed to describe the phosphorus trends existing in the A.S.F.T.L. data.

The complete regression technique is difficult to apply efficiently in a computer program to the A.S.F.T.L. data for the whole province. It would, however, be more applicable to smaller areas and should produce some interesting trends that can be numerically defined. Response surfaces for the regional trend developed by averaging techniques such as the circle method are not numerically defined.

(g) Multiple regression analysis - - A step-wise multiple regression program was applied to the data in an attempt to formulate a mathematical relationship of average phosphorus per township (dependant variable) with township and range values acting as independent variables. Work by Walker et al. (op. cit.) indicated that simple polynomials, up to cubic terms would be appropriate for the analysis of soil trends. The results, however, indicated that very little relationship exists. The F-ratio for the 9 independent variables - - 2 linear, 3 quadratic, and 4 cubic terms - - and 305 observations (township averages) was 1.05 which is not significant at the 5% level. A low  $R^2$  value of .031 was obtained with a standard error of 19.0 (with a  $\bar{Y}$  mean of 44.8). Several runs were made involving higher order polynomials and similar results were obtained. The excessively low  $R^2$  values coupled with the almost impossible task of applying such a technique to the mapping of soil test data on a large scale prevented any further inquiry into the technique.



In summary, then, a number of different mapping techniques were applied to the soil test data using a selected trial area. The results show that the "expected value" method, the orthogonal polynomial methods, and the regression method are not suitable for the mapping of such data; although other authors and researchers have found these techniques to be successful for their particular data. The method of "eyeballing" can be used with the aid of a computer to sum, average and round-off the results. Experience is the criteria for such a technique. This leaves the profile and circle methods, both very applicable to the data. Of these the circle method has been selected mainly because of its completeness and accuracy. In addition, a similar technique presented by Bayrock and Pawluk (op. cit.) has already been used and proved successful for mapping on a provincial basis.

Alberta soil nutrient maps: The details of the computer technique used to develop the Alberta soil nutrient and pH maps from the A.S.F.T.L. data are discussed in Appendix III. A listing of the Fortran IV program and a written description is available in the Department of Soil Science, University of Alberta.

The following will be a discussion of the results obtained from the computer mapping of the data. But first, a brief description of the computer program is necessary to clarify some obvious ambiguities that might be questioned in comparing the final plotted maps.

The program used is fairly complicated. There are 44 variables that can be used to control the final output of the map. Most of these are discussed (and in some cases their results illustrated) in Appendix III. There are five major categories of control. Firstly, there are the time





period and crop selection controls which have usually been set on the basis of soil area results for the nutrient similarities under different crops. Secondly, there are weighting controls based on the number of cores that were used to make up a sample. These controls can be used to weight each of the individual samples or they can be ignored. Thirdly, there are weighting controls based on the number of samples representing a township. These can be used to assign a weight to each township, calculate an average for each township, or they can be ignored. Fourthly, there are the circle controls which determine the size and weighting of the mapping circle, the size of the critical circle, and the critical number of samples required before a map calculation is made. Fifthly, there are the map size control and rounding factors which can be varied.

An adjustment in any or some of these controls will most likely result in a slightly different map. For example, maps covering different time periods with different controls used for each map sometimes show different results. This leads to a question as to the cause of this difference -- is it the different time period, or is it the different map control variables that cause this difference? The answer is contained mainly in Appendix III.

In summary, however, resulting maps show these trends. Both the weightings based on the number of cores per sample and the number of samples per township generally have very little effect on the final map values (for the maps reproduced in this section). Thus, differences between sampling years, if different control variables are used, can largely be attributed to real differences between years; not map control variables. The weightings that do make a difference are those for the concentric





circles within the mapping circle, in effect, they are the principal control variables. However, unless otherwise stated, all weightings used are 4, 4, 3, 3, 2, 2, 1, 1 for the eight concentric circles, from the inner circle outward, respectively.

One other noticeable difference between the maps that will be presented is the extent of the mapped areas relative to that of the areas for which no value is plotted. The magnitude of these areas will vary from map to map. No map value is plotted unless there are a certain number of samples within the critical circle; the value of this critical number can be varied. From Tables XIII and XIV, presented in the previous section on sampling, one can see that sampling numbers will vary considerably within the critical circle depending upon the time period and the crop types selected. Therefore, because different time periods and crops have been selected for the mapping of various nutrients, slightly different sampling patterns will appear on the different maps. Another reason, and perhaps the main reason, is the selection of the critical number. This is a critical selection that has to be based on the results of the above mentioned tables as well as the following: the size of the critical circle; the assigned sample and township weights (if any); the degree of accuracy required and, closely related, the desired completeness of the map; and finally the results that similar trial runs have rendered. The selection is sometimes quite difficult, especially where sample and township weighting controls have been used. If the critical number is set too high, the desired accuracy may well be obtained, but only those areas with this accuracy will be displayed, the end result being a patchy map with large open and disjointed segments. A critical number set too low, on the other hand, will



decrease accuracy, often produce erratic "edge" effects, and display spotty localities (where very few samples exist) in large open areas. Thus a compromise often had to be made between desired accuracy and completeness; and the latter was heavily weighted where short time periods and few cultural practices were selected, as would be the case of nitrogen on fallow for one sampling year.

The size of the plotted map was kept constant at 1 inch to every 70 miles. All maps are presented exactly as constructed by the computer and plotted by the accompanying plotter, but with the top boundary manually removed to meet page size requirements. For the purpose of locating contours or individual values for any township, an exact plastic overlay is available in a pocket attached to the back cover of this thesis. The overlay contains familiar landmarks and labeled townships and ranges, enabling fast and precise location of any area. Approximate legal locations will be referred to from time to time in describing the features of the individual maps.

The first part of the discussion will deal with the distribution of sample numbers within the province followed by some of the mapped results of nitrogen, on both fallow and stubble, phosphorus, potassium, and pH. A complete discussion of the results illustrated in each map would be an enormous task and no attempt will be made to undertake such a project. The development of the computer programs for future use and a presentation of major results in a published text for future reference were felt by the author to be the areas of primary concern in this study. However, a brief, general description of each map will be presented with some of the major trends noted.



(a) Sampling density - - An accurate indication of areas of high and low sample numbers is necessary in complete interpretation of the results from the mapping program. Such an indication has been previously presented for the county and districts in Figure 28.

By a slight alteration of the Alberta mapping program, the average number of samples per township could be plotted. Figure 39 is a map of Alberta showing coded numbers representing the number of samples available in each township for the A.S.F.T.L. data from April, 1962 to March, 1969 with all months and crops included. Townships with no samples available have no value plotted. From such a map, broad areas of heavy and light sampling density can roughly be distinguished by "eyeballing". However, by mapping the number of samples, a more refined distinction can be made. Such a map enables one to distinguish accurately the areas with high and low sample numbers. Figure 40 is a contoured map of Alberta showing areas of high and low sample concentration. The Edmonton and Camrose regions appear to have the heaviest sampling with an average of over 55 samples per township, followed closely by the areas surrounding Red Deer, Calgary, and Lethbridge with averages around 46 to 55 samples per township. The area south of Consort, through Medicine Hat to the border, appears to be a very poorly sampled area with under 10 samples per township. The Peace River area is not as densely sampled as the rest of the province appears to be, the Fort Vermilion area being poorly sampled. In general, the results show similar trends to those discussed in the county maps.

(b) Nitrogen - - Figures 41 and 42 are maps of Alberta showing coded numbers for relative levels of nitrate nitrogen in every township with 2 or more samples available. Figure 41 shows the levels of nitrate nitrogen on





Figure 39. Map of Alberta showing the average number of samples per township. Time period: Apr., 1962-Mar., 1969. All crops.

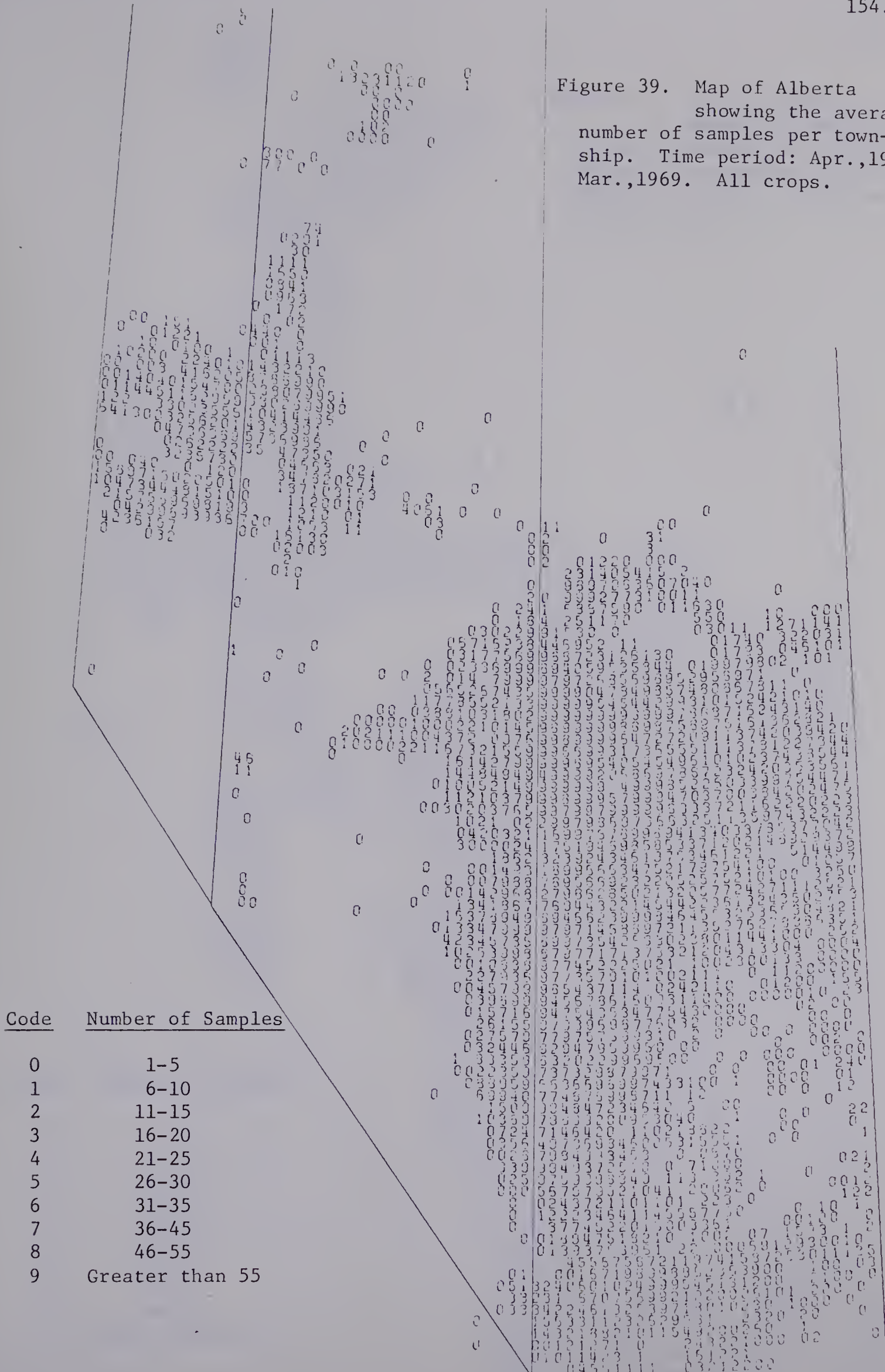






Figure 40. Map of Alberta  
showing distribution  
of sample numbers on all crops  
for the time period: Apr.,  
1962-Mar., 1969. Circle weights:  
4 4 3 3 2 2 1 1.

Code      Number of Samples

0	1-5
1	6-10
2	11-15
3	16-20
4	21-25
5	26-30
6	31-35
7	36-45
8	46-55
9	Greater than 55

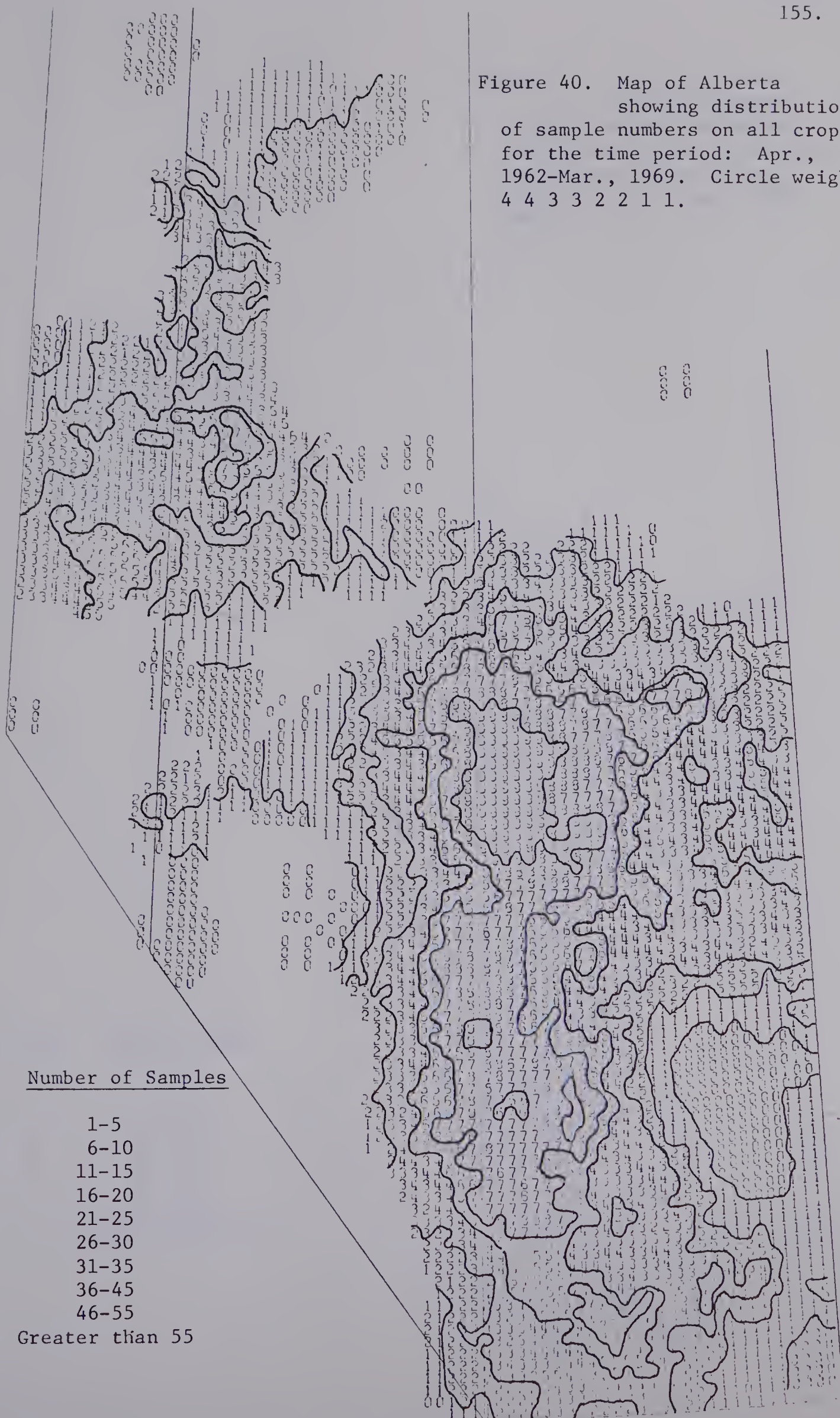




Figure 41. Map of Alberta showing nitrate nitrogen values on fallow for townships with 2 or more samples available. Time period: Sept. - Apr., 1962-1969.

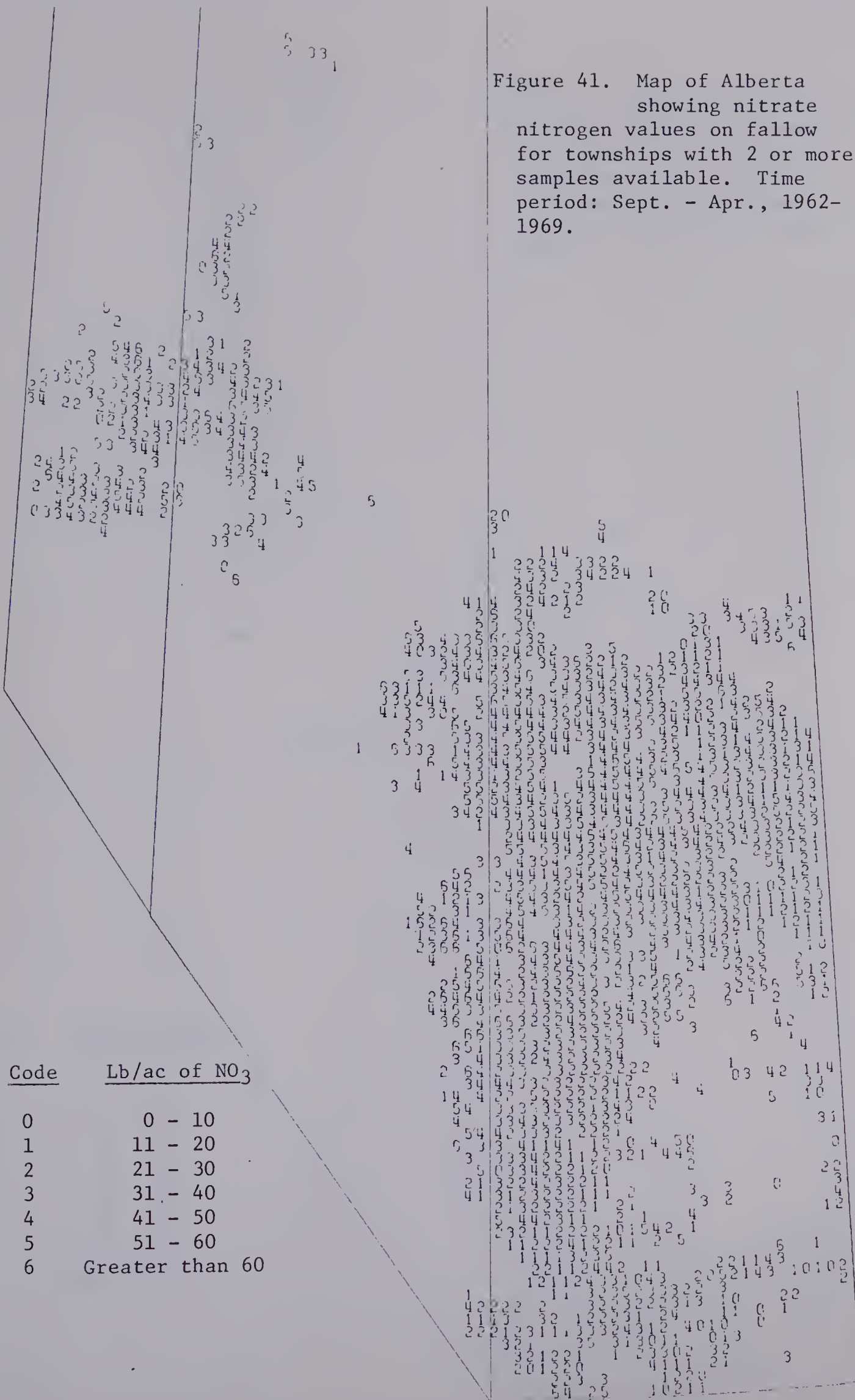
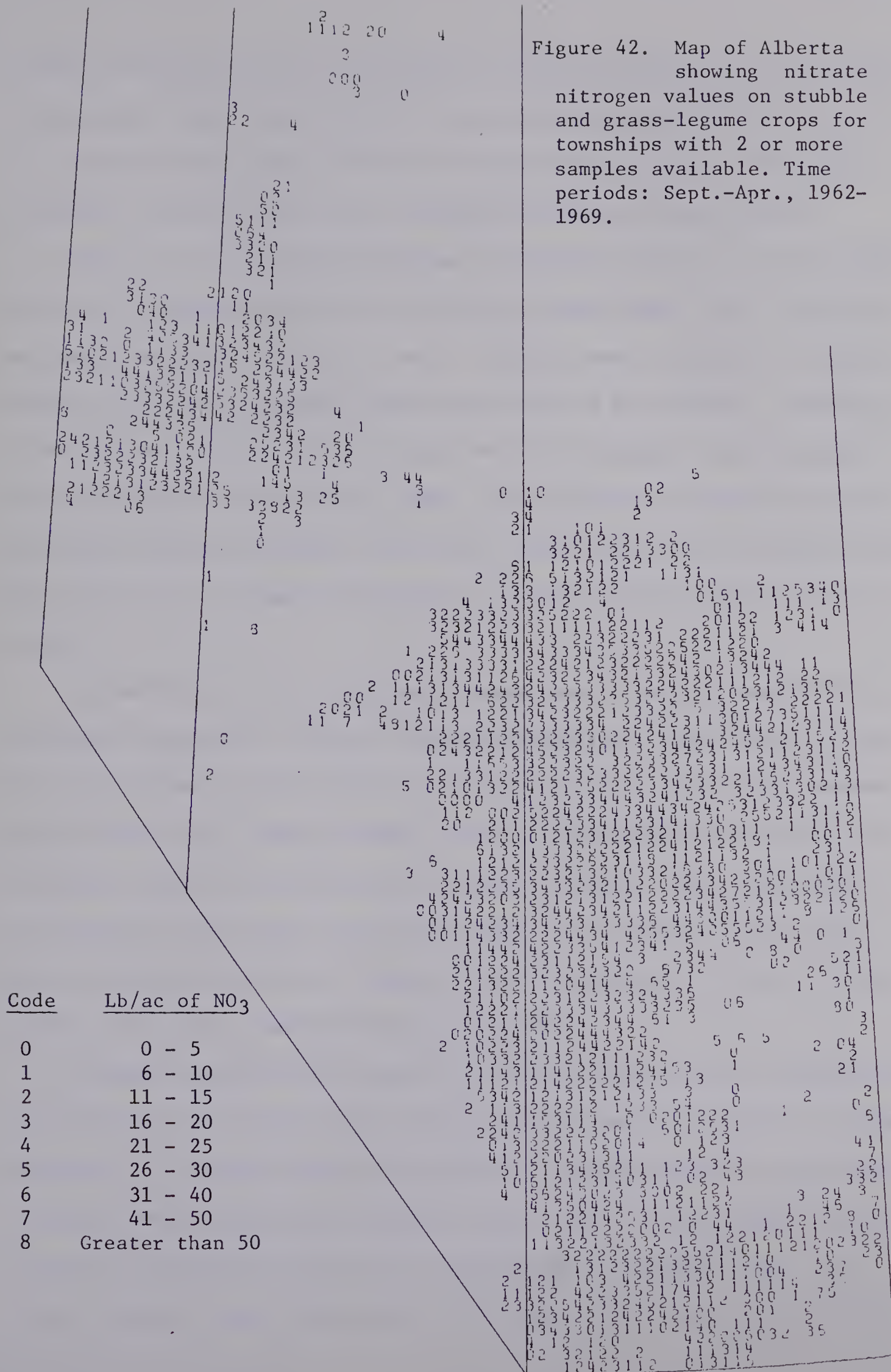






Figure 42. Map of Alberta showing nitrate nitrogen values on stubble and grass-legume crops for townships with 2 or more samples available. Time periods: Sept.-Apr., 1962-1969.





fallow and Figure 42 shows the levels of nitrate nitrogen on stubble and legume-grass crops. Both maps have been plotted using all the soil test data available since April, 1962 up to March, 1969, but only the months September through to April were included for any one sampling year.

There are two major disadvantages of results presented in this manner, and this is particularly true for nitrate nitrogen where large variations are known to occur from year to year. Firstly, such a display of numbers makes it difficult to depict trends that exist in the results. Secondly, averages over the years do not ensure that all townships contain samples from every year or even several years. Some township averages may be predominantly biased by samples from certain years and others from other years. For this reason, nitrate nitrogen data is better displayed for one year at a time.

In the light of this knowledge a comparison of the two maps should be conducted cautiously. The first visual observation is that stubble crops are more predominant than fallow, as indicated by the larger gaps present in the fallow map. Fallow, however, has the higher values of nitrate. No clear-cut trends in the data can be readily surmised. Both maps do give an indication that nitrate levels are higher in the central part of the province than the southern. However, this is especially difficult to perceive on the stubble and grass-legume crops map.

A mapped summary of the data over the years does have the advantage of spotlighting those regions where low and high values are inclined to appear frequently. The application of the circular mapping function to the data definitely helps off-set any specific yearly biases. The mapping function takes into account the data over a large number of townships and, as a result, usually samples from most of the years. Thus, such a map would be.





valuable in defining areas of high and low nitrate values, although no conception of the yearly variation can be ascertained.

Figure 43 is a map of Alberta showing the distribution of nitrate nitrogen on fallow with all the available soil test results considered. The map shows that Edmonton, Red Deer, and Calgary areas generally have the largest accumulation of nitrate with averages between 41 and 50 lb/ac. There is a trend for these high values to fall in the Black soil zone. The southern part of the province appears to show the lowest accumulation with 21 to 30 lb/ac. The Peace River region is variable, but with the majority of soils having 31 to 40 lb/ac. It is interesting to note that the patch of Gray Wooded soils in the vicinity of Cooking Lake is easily distinguished on the map. Trends in the Dark Gray and Gray Wooded soils to the north and northeast of Edmonton are variable, partly because of the sparse sampling.

Figure 44 is a map of Alberta showing the distribution of nitrate nitrogen levels on stubble and grass-legume crops for the soil test results that have been collected since 1962. The map shows a seemingly more complicated surface than the fallow map. This, however, is mainly due to the different rounding factors used to derive the number codes, and should be taken into consideration in any comparison. The trends in the map are difficult to relate to any particular soil areas. The majority of the regions show an average of around 11 to 15 lb/ac. Extremes in the mapped values for nitrate levels after stubble and on grass-legume crops do not usually appear, and if they do it is usually in areas where few samples have been obtained, such as that area north of Medicine Hat.

The area summaries have previously emphasized the large variation that nitrate nitrogen exhibits from year to year. Figures 45, 47, 50, and 52



Figure 43. Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: all months, 1962-1969. Circle weights: 4 4 3 3 2 2 1 1.

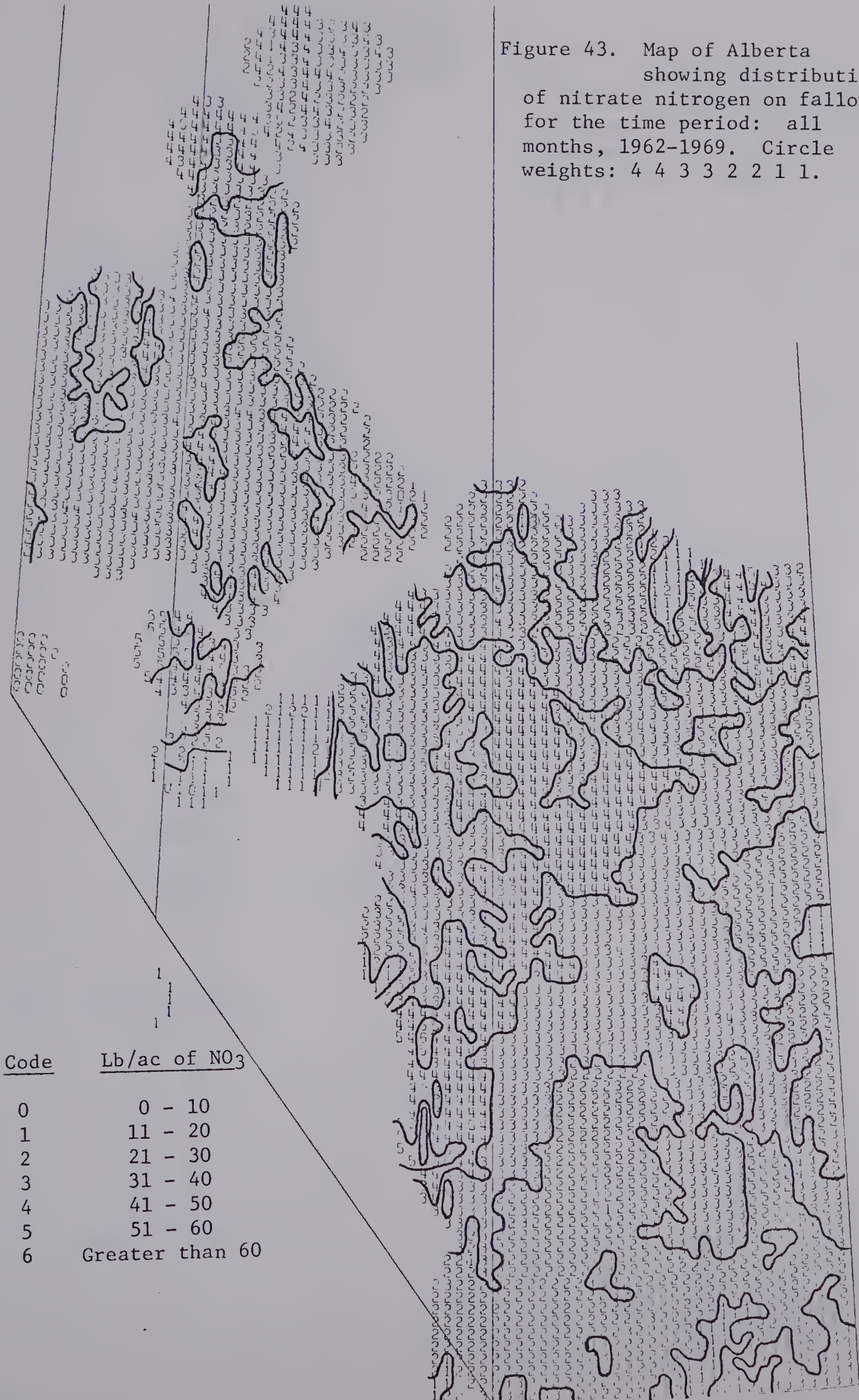
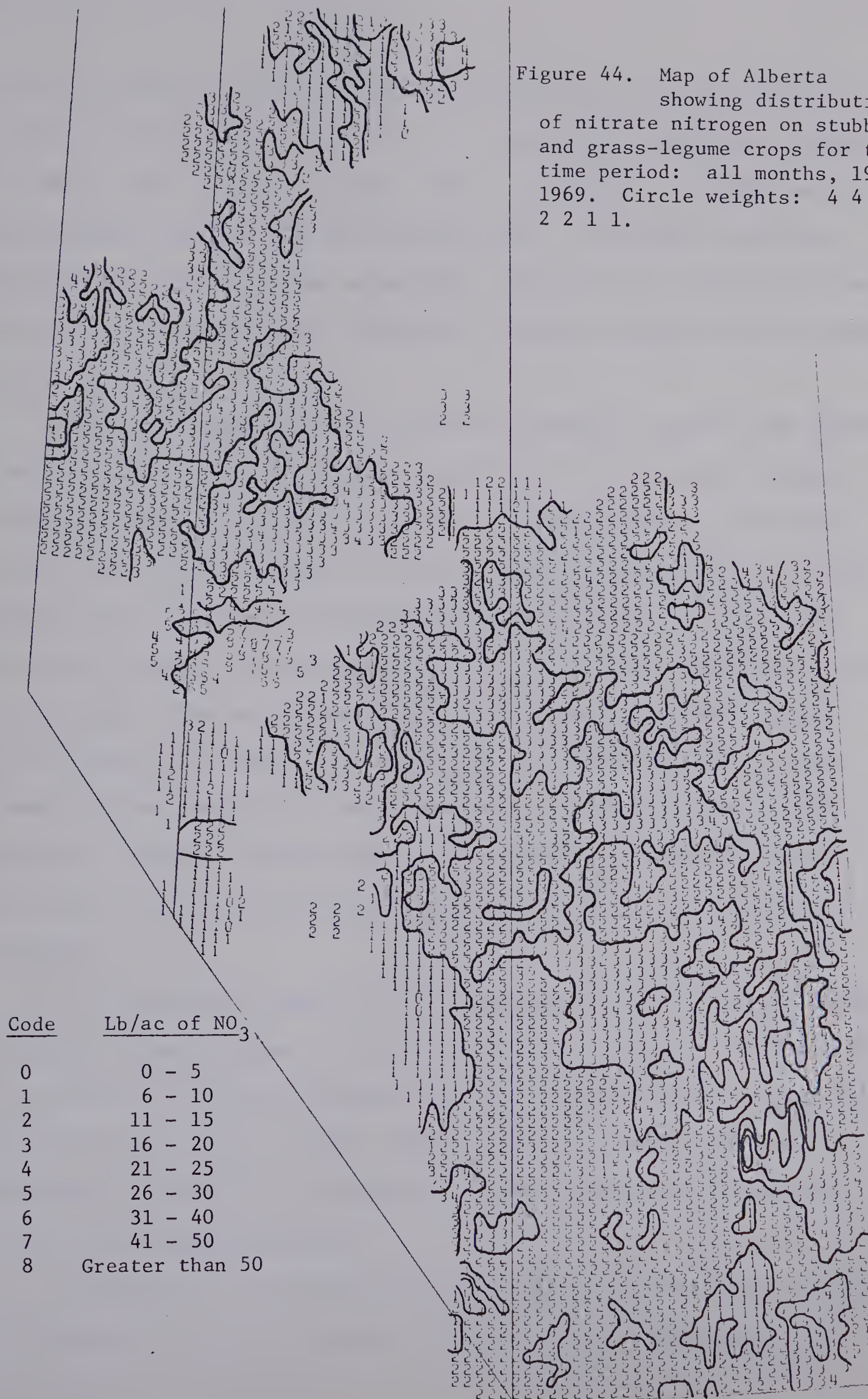






Figure 44. Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: all months, 1962-1969. Circle weights: 4 4 3 3 2 2 1 1.







show the distribution map for nitrate nitrogen for the sampling years 1968-69, 1967-68, 1966-67, and 1965-66, respectively. Similarly, Figures 46, 48, 51, and 53 show the mapped values for nitrogen on stubble and grass-legume crops for the same sampling years. The months covered by each sampling year are shown on each map. Figure 49 is a similar plot to Figure 48, the sampling year 1967-68, but a larger rounding factor has been used.

A general comparison between fallow and stubble crops for each sampling year shows that these two cultural practices do not necessarily respond similarly from year to year. For example, the map for the sampling year 1968-69 shows nitrate values on fallow to the west of Edmonton to be quite high relative to those in the southeastern part of the province. This same trend, however, is not present in the results from stubble and grass-legume crops. However, in some years and locations, a matched response can be noted. For example, in the sampling year 1967-68, the area about township 30 and range 15, W4 shows the same pattern of response on both practices. Similar inconsistencies in the relative nitrate fluctuations on fallow and stubble crops have been previously pointed out in the area summaries.

As illustrated in Figure 48, the sampling year 1967-68 shows a fairly detailed contouring scheme for nitrate on stubble. Often, for such a map, it is easier to distinguish general trends by making larger intervals into which the results can be coded. This has been done and the results are displayed in Figure 29. The map shows a much smoother surface on which trends can be rapidly depicted.

A comparison of the nitrate levels over different years shows extreme differences in not only the general levels of nitrate, but location of



Figure 45. Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1968-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

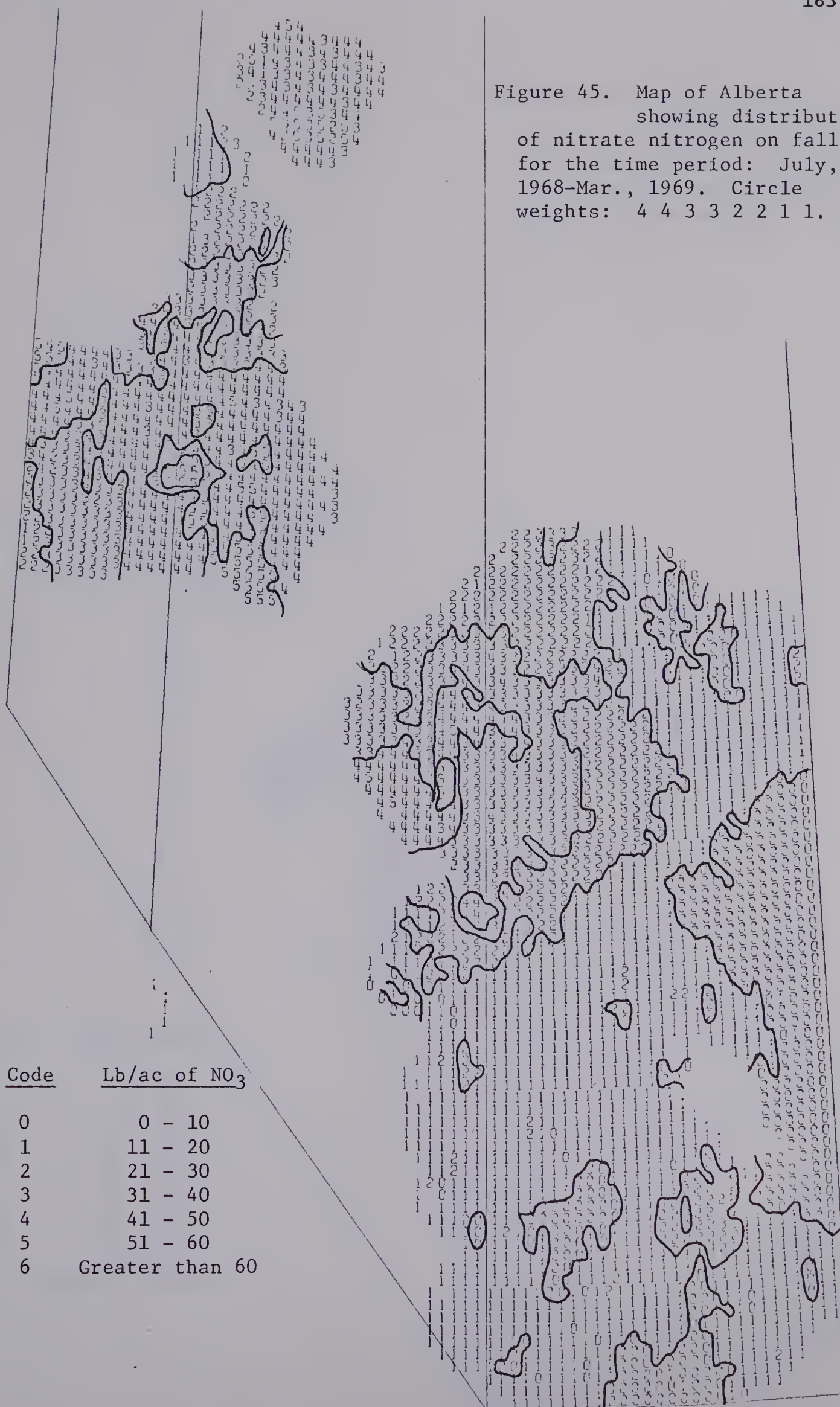




Figure 46. Map of Alberta  
showing distribution  
of nitrate nitrogen on stubble  
and grass-legume crops for the  
time period: Sept., 1968-Mar.,  
1969. Circle weights: 4 4  
3 3 2 2 1 1.

Code	Lb/ac of $\text{NO}_3$
0	0 - 5
1	6 - 10
2	11 - 15
3	16 - 20
4	21 - 25
5	26 - 30
6	31 - 40
7	41 - 50
8	Greater than 50

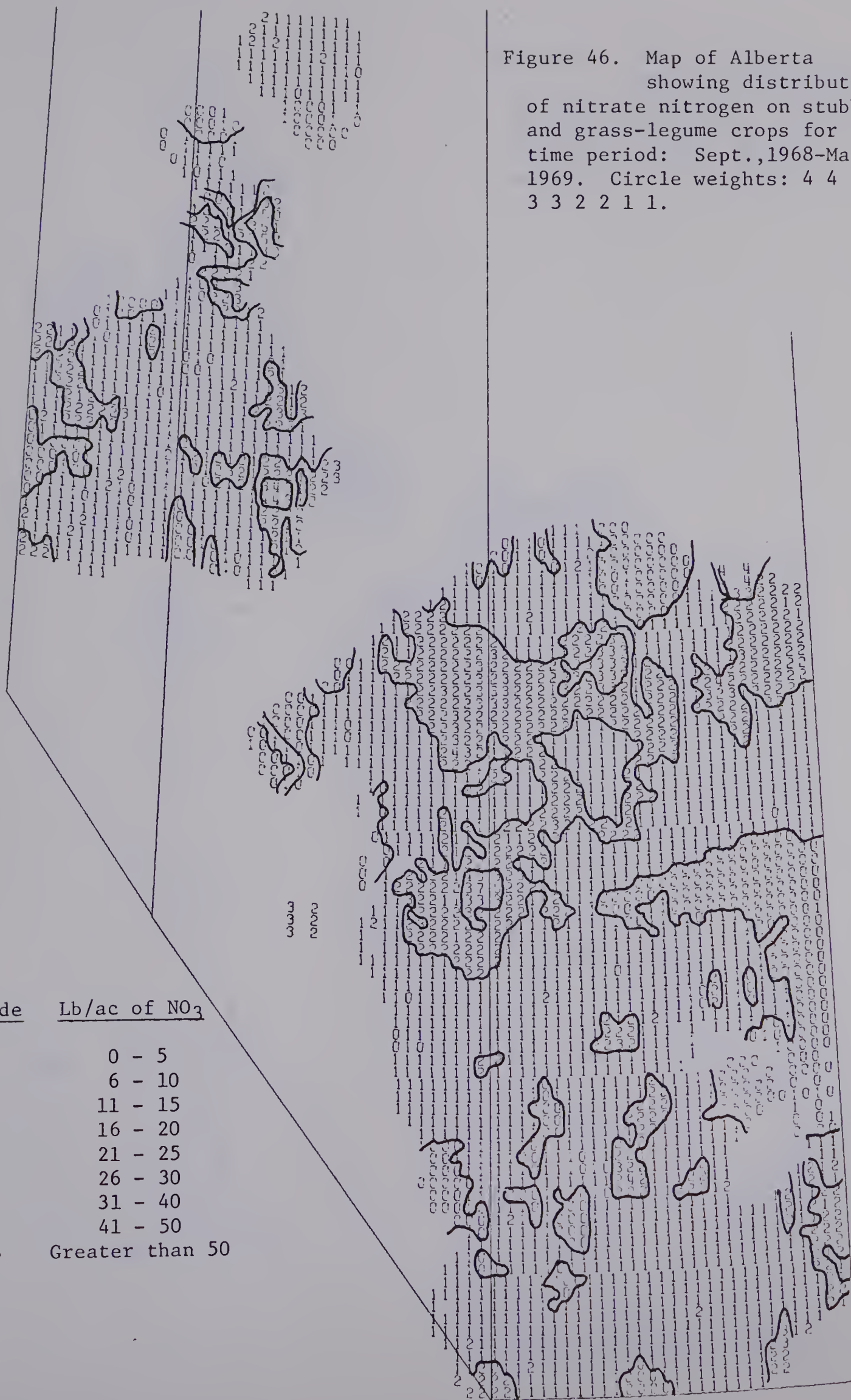






Figure 47. Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1967-June, 1968. Circle weights: 4 4 3 3 2 2 1 1.

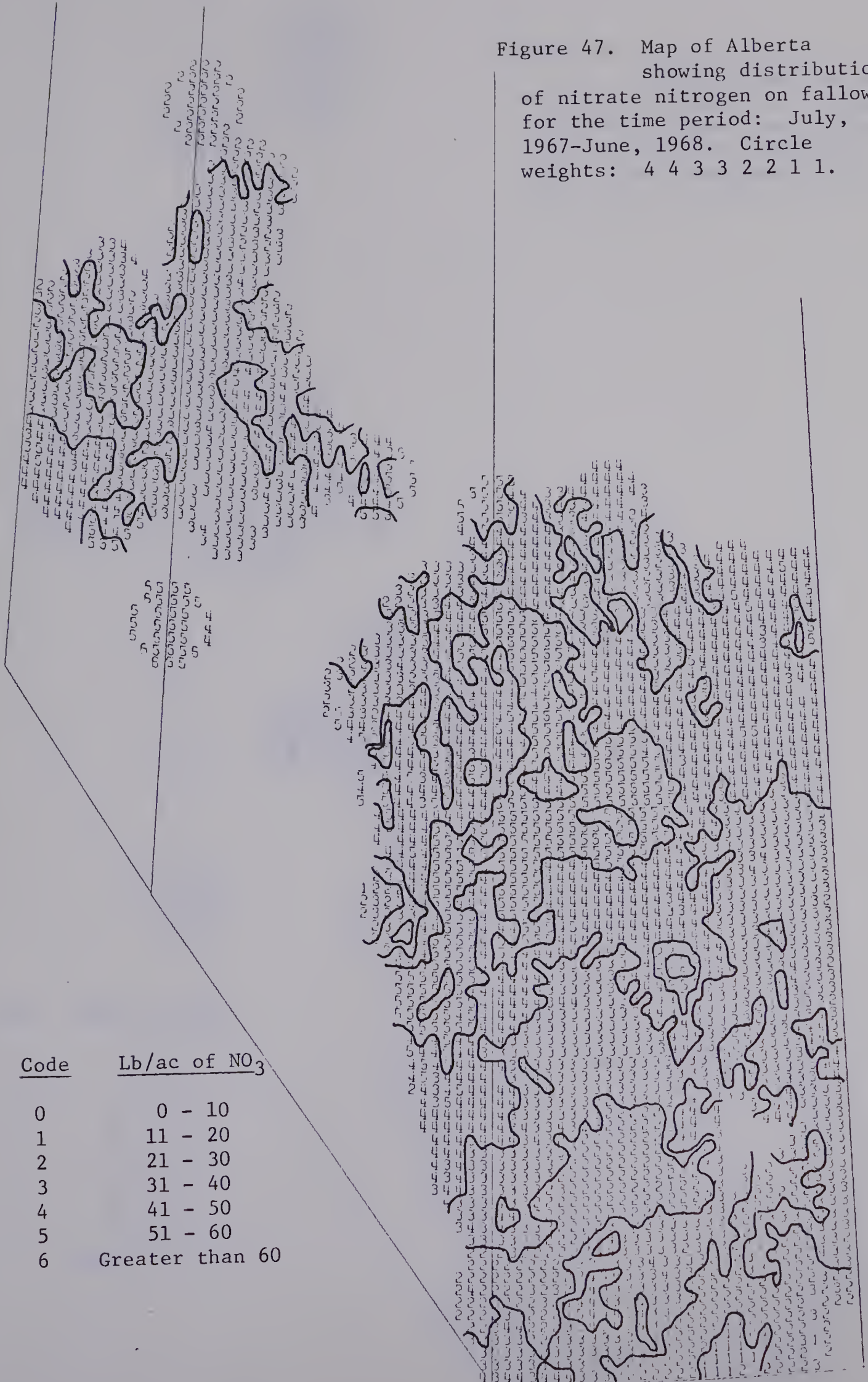






Figure 48. Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: Sept., 1967-Apr., 1968. Circle weights: 4 4 3 3 2 2 1 1.

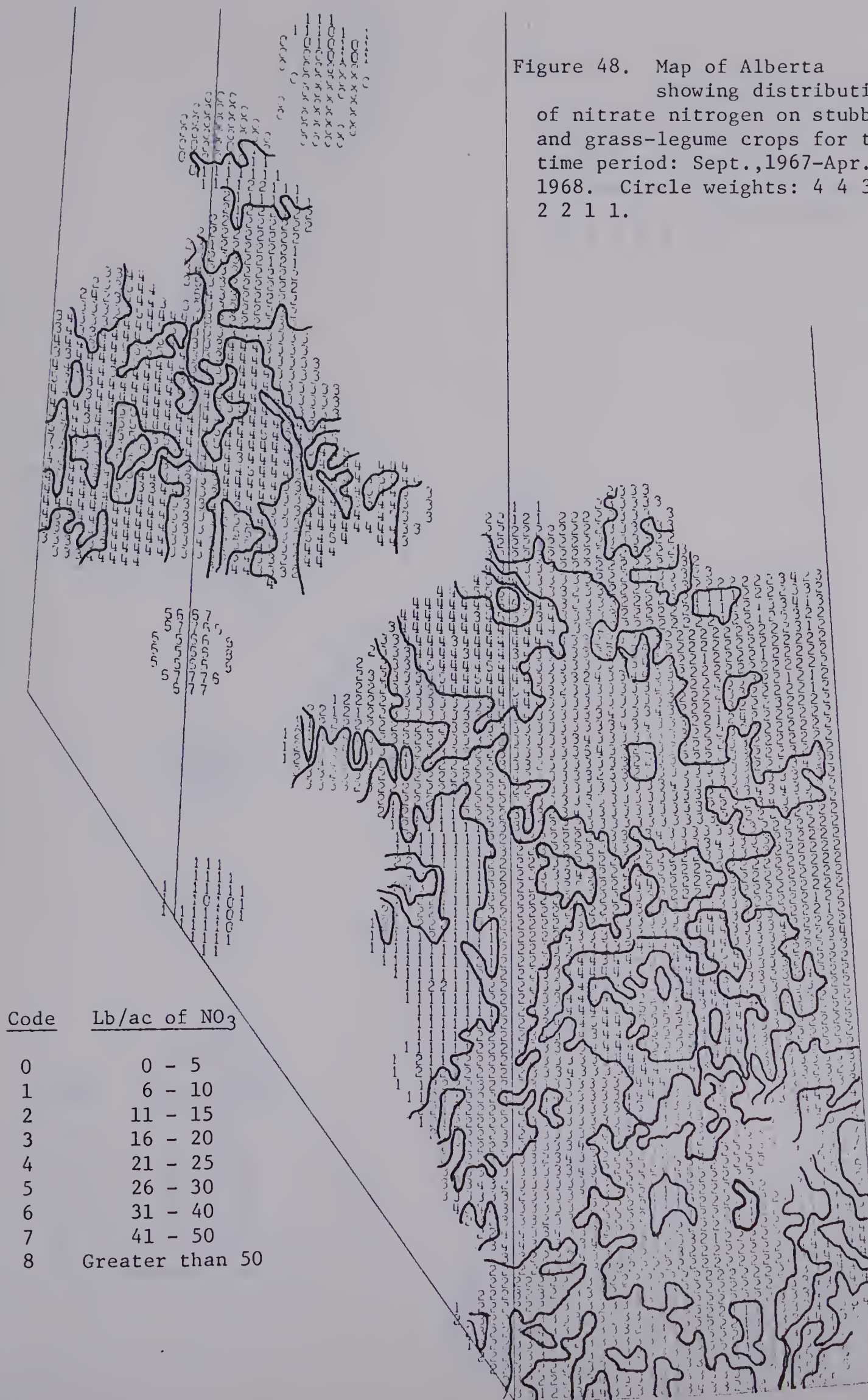




Figure 49. Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops, using larger rounding factors, for the time period: July, 1967-June, 1968. Circle weights: 4 4 3 3 2 2 1 1.

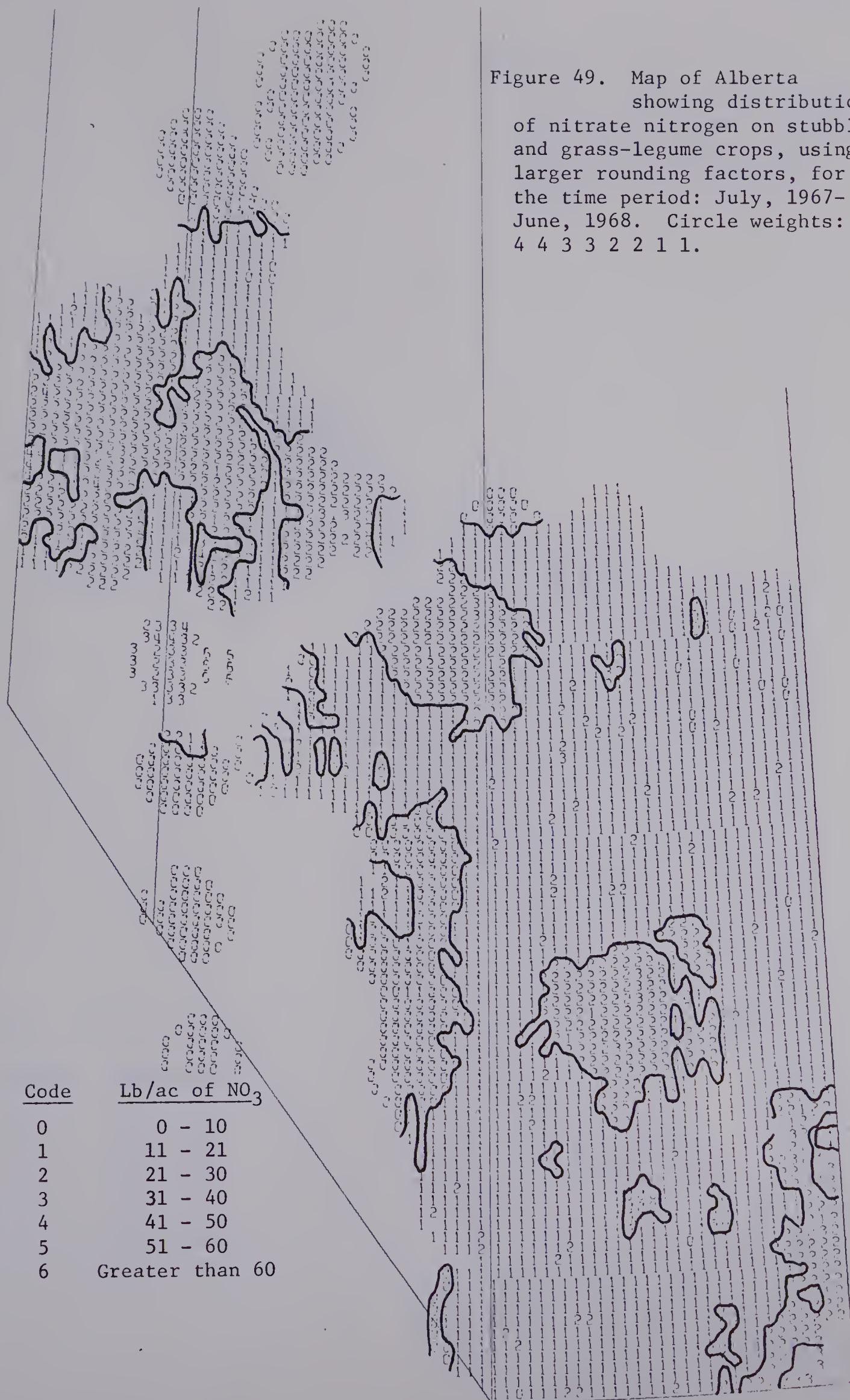






Figure 50. Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1966-June, 1967. Circle weights: 4 4 3 3 2 2 1 1.

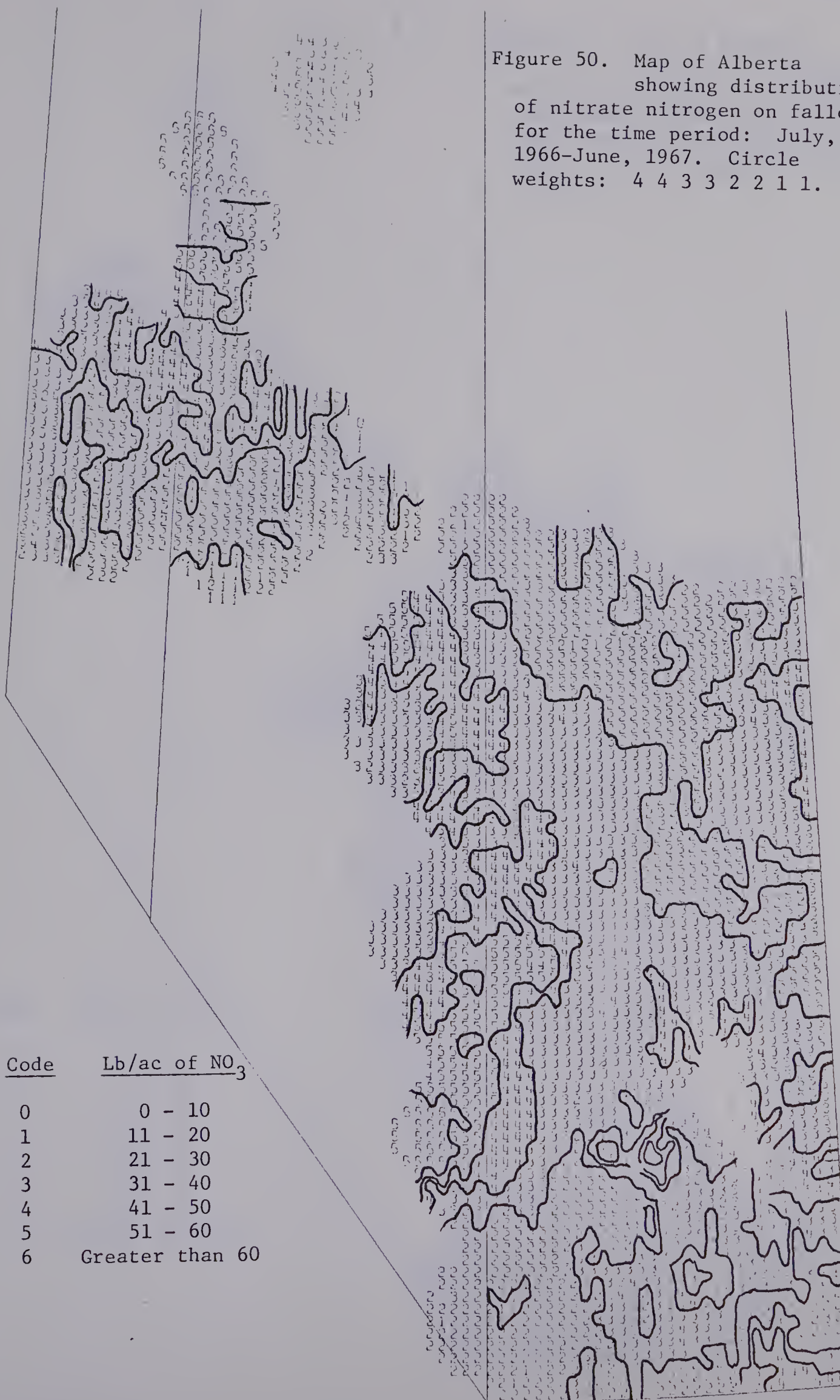






Figure 51. Map of Alberta showing distribution of nitrate nitrogen on stubble and legume grass-crops for the time period: Sept., 1966-Apr., 1967. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of NO <sub>3</sub>
0	0 - 5
1	6 - 10
2	11 - 15
3	16 - 20
4	21 - 25
5	26 - 30
6	31 - 40
7	41 - 50
8	Greater than 50

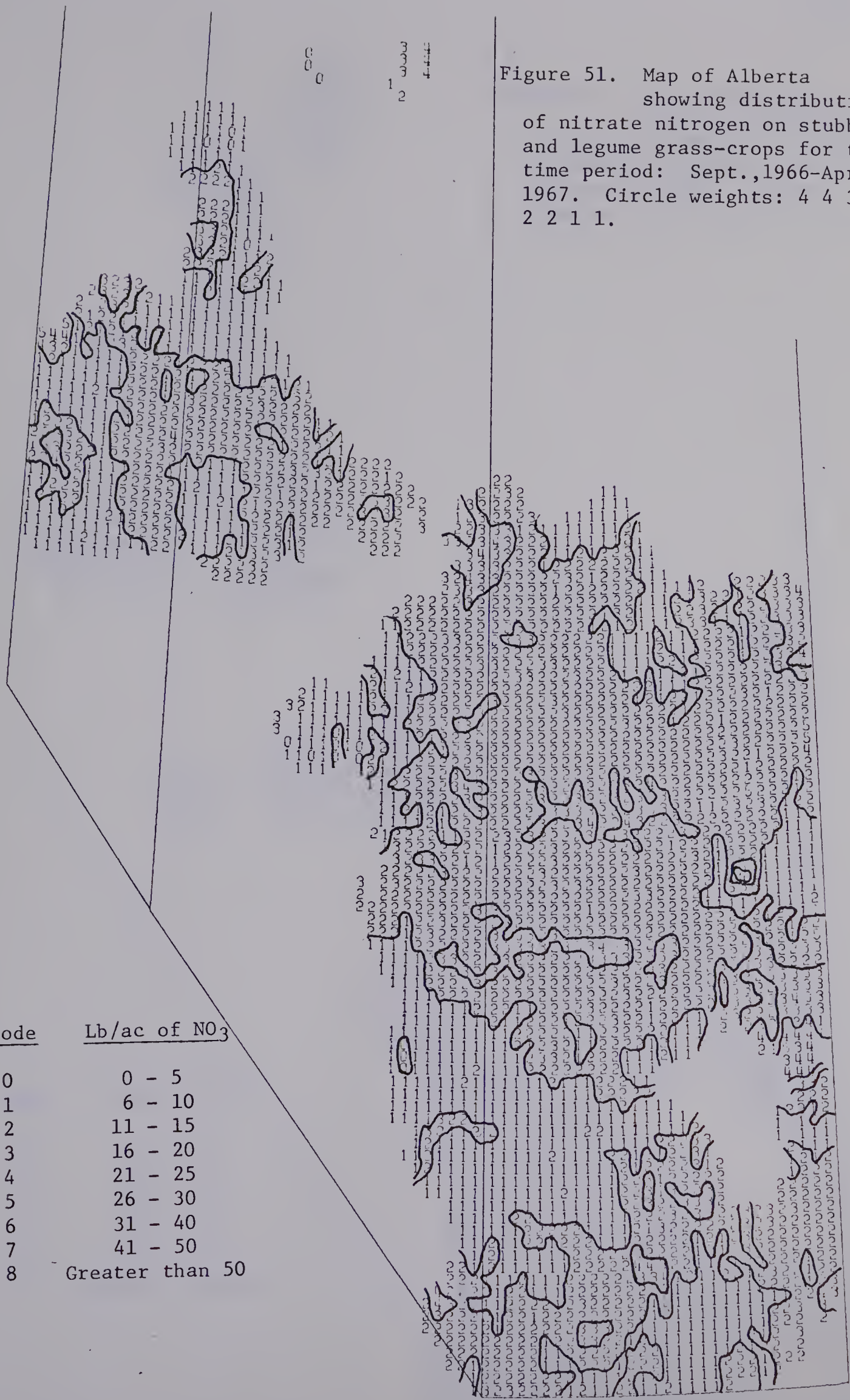




Figure 52. Map of Alberta showing distribution of nitrate nitrogen on fallow for the time period: July, 1965-June, 1966. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of NO <sub>3</sub>
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 60
6	Greater than 60

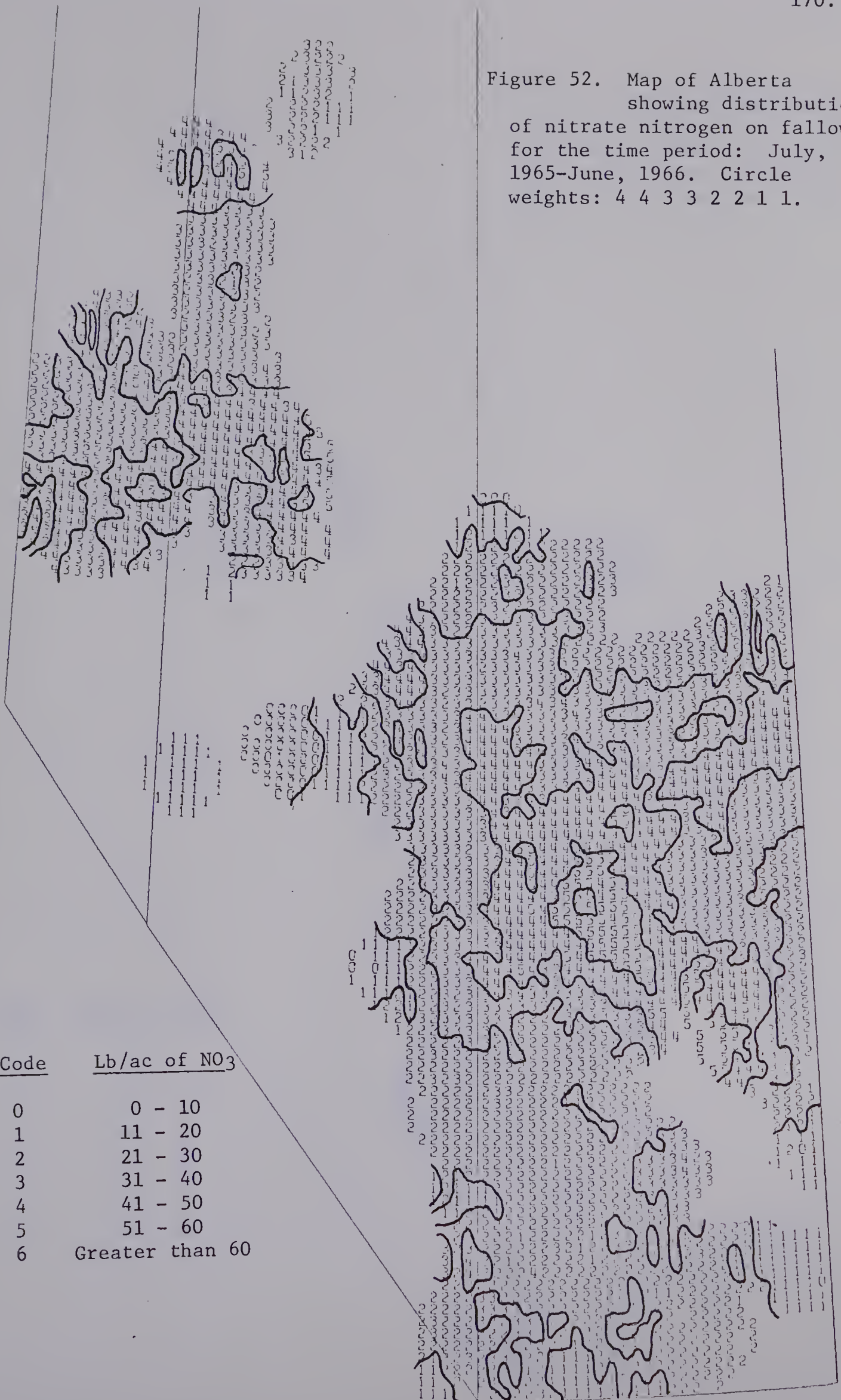
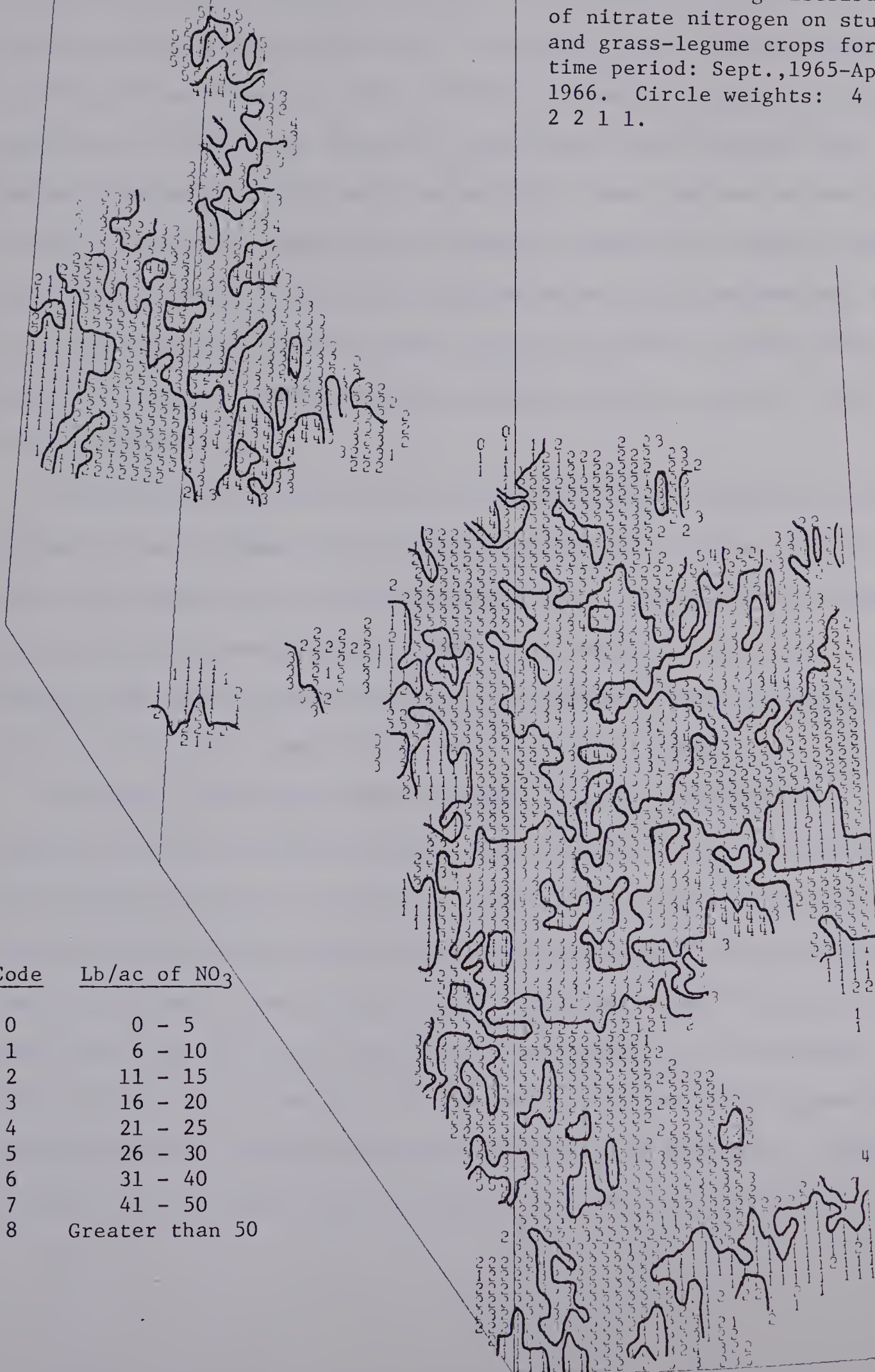






Figure 53. Map of Alberta showing distribution of nitrate nitrogen on stubble and grass-legume crops for the time period: Sept., 1965-Apr., 1966. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of $\text{NO}_3$
0	0 - 5
1	6 - 10
2	11 - 15
3	16 - 20
4	21 - 25
5	26 - 30
6	31 - 40
7	41 - 50
8	Greater than 50





similar high and low areas for each year. The sampling year 1968-69 appears to be a very low year, while the sampling year of 1967-68 shows very high accumulations of nitrate on fallow. None of the years show the same pattern of variation, in fact, a comparison of all four maps displaying nitrogen on fallow, makes any similar trends from year to year hard to depict. On fallow, generally, the Edmonton area shows a higher accumulation than the surrounding areas, with the areas in the southeastern part of the province exhibiting the lowest levels of nitrate on fallow. The Peace River region, although variable, usually follows a similar trend to the Edmonton area.

A comparison of nitrate levels on stubble and grass-legume crops over the years shows no distinctly similar trends for any two years. As on fallow, the sampling year 1968-69 shows the lowest accumulation of nitrate with the preceding sampling year showing the highest accumulation. Any definite high or low trends for particular areas are hard to distinguish because of the almost random fluctuation from year to year.

In summary, then, maps made to compare nitrate nitrogen levels indicate a number of general trends. Fallow levels of nitrate are higher than stubble in any year for any area. Fallow and stubble fields do not always respond similarly, with fallow fields showing the greatest variation from year to year. Areas that are continually high or low from year to year relative to the rest of the province are difficult to distinguish unless the data are mapped over all the years. Nitrate levels appear to be closely related to weather conditions, moreso than to soil areas, although the final value is most likely an interaction between the two factors.





(c) Phosphorus - - A map of Alberta showing coded numbers for the average phosphorus value per township on fallow, stubble and grass-legume crops for all the soil test data from April, 1962 to March, 1969 is shown in Figure 54. No value has been plotted on the map, unless two or more samples existed in the township. The crop groupings chosen to summarize the data in map form, have been selected on the basis of the crop comparisons for phosphorus levels previously completed in the area summaries. Such a presentation allows one to examine individual values per township, but does not permit easy visualization of any apparent trends.

Figure 55, on the other hand, permits the easy visualization of trends. Figure 55 is a map of Alberta showing the phosphorus distribution on fallow, stubble, and grass-legume crops for all years from April, 1962 to March, 1969. The contoured map points out district trends. The southwestern part of the province around Pincher Creek displays a fairly low level of phosphorus with averages around 11 to 20 lb/ac. The area extending from Calgary southeast to Medicine Hat shows an increase with 21 to 30 lbs/ac on the average. The areas to the northwest and northeast of Edmonton are quite high in phosphorus, usually around 41-60 lb/ac. A distinct island of coded 4's, surrounded by coded 3's, appears to the southeast of Edmonton. This island approximately outlines a group of Solonetzic soils in that area. The Peace River region appears to have similar levels of phosphorus to those about the Edmonton area. The area around Grande Prairie and to the southwest shows a lower content of phosphorus than the neighboring areas to the northeast. The Athabasca region appears to have the highest levels of phosphorus with averages usually between 51 to 70 lb/ac.

Figure 56 portrays a more general and easier visualization of the trends.



Code

Lb/ac of P

0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

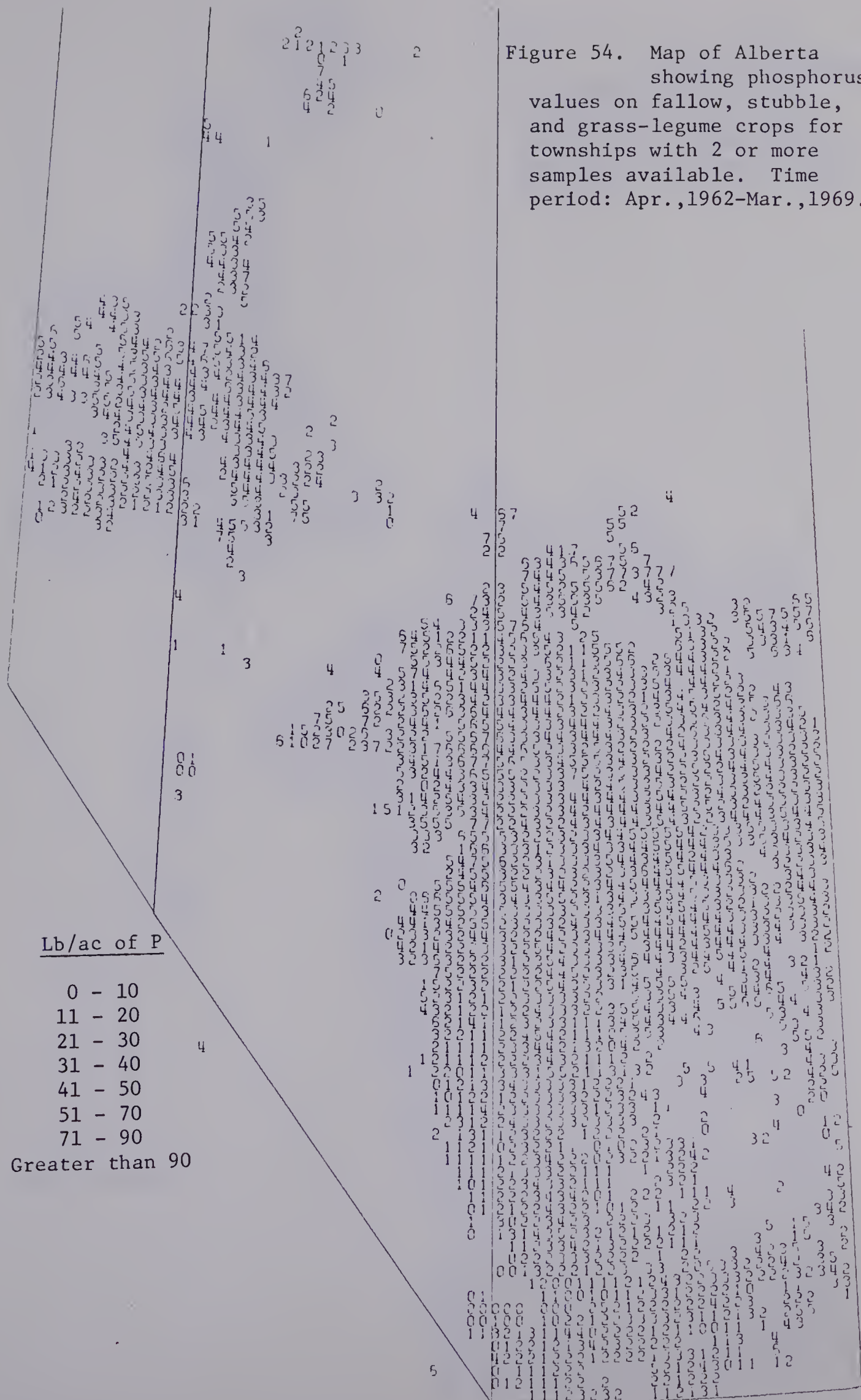






Figure 55. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: Apr., 1962-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

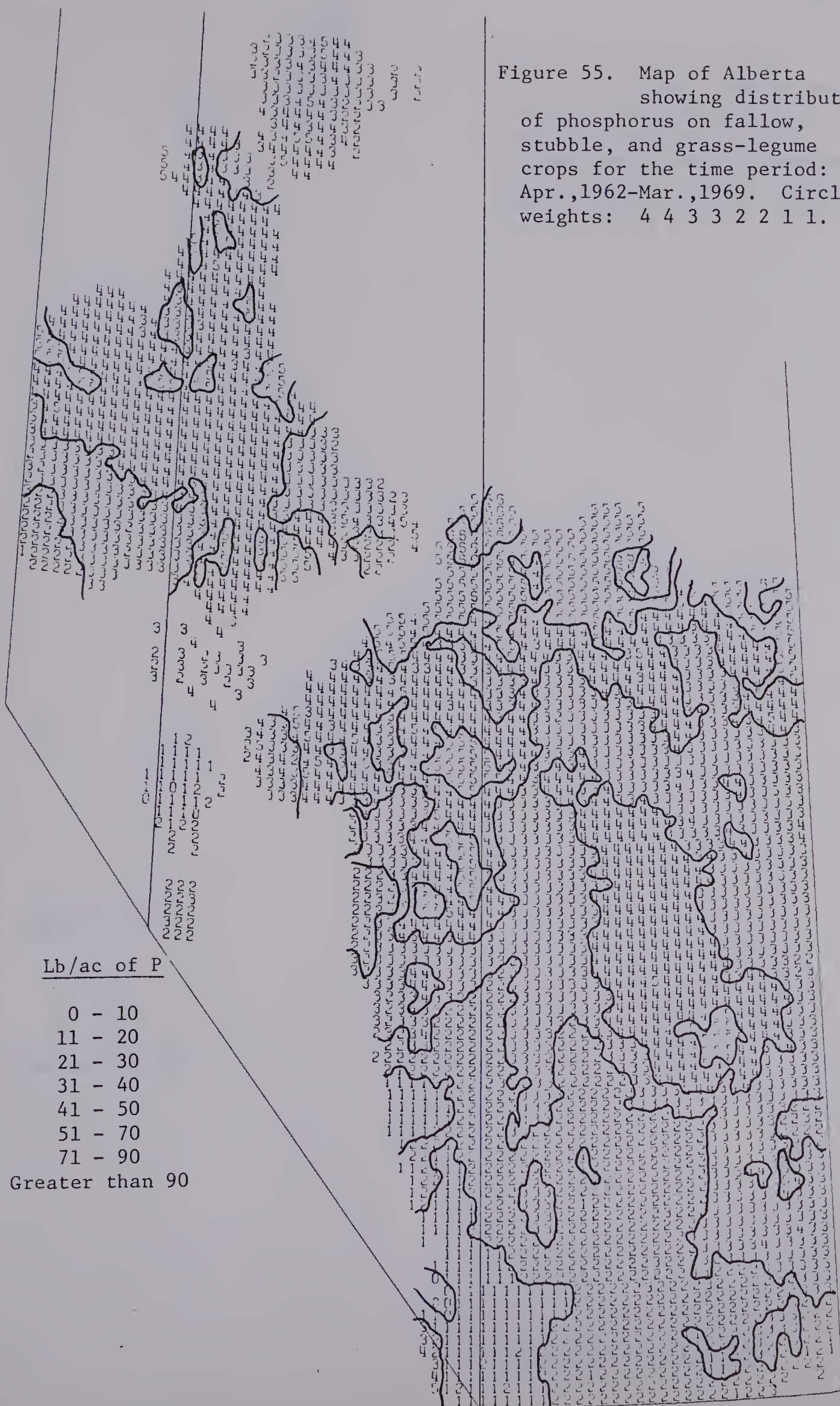






Figure 56. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops, using larger rounding factors, for the time period: Apr., 1962-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 20
1	21 - 40
2	41 - 60
3	61 - 80
4	Greater than 80

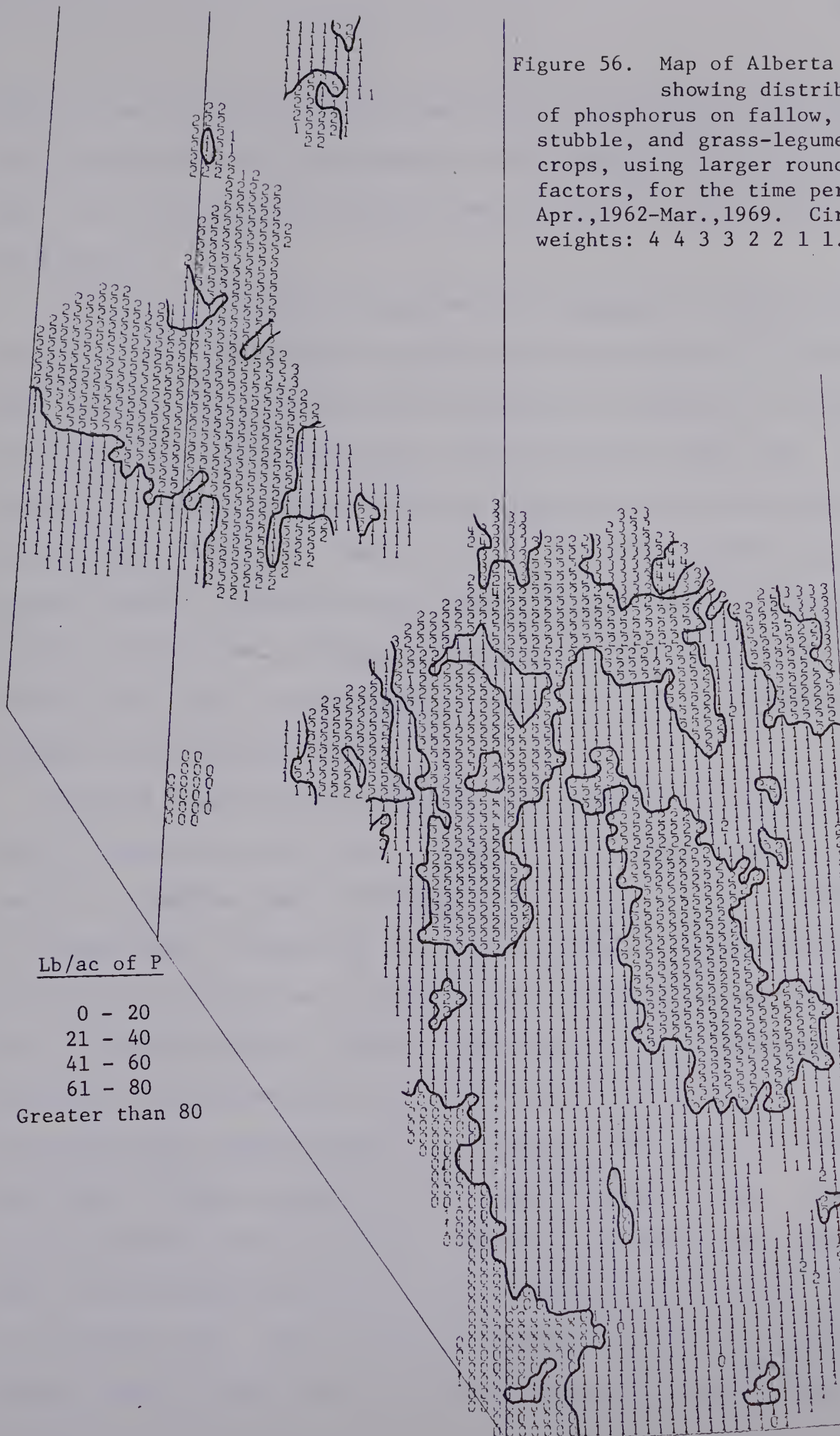




Figure 56 has the same time period and crops as the previous figure, but the rounding factor has been altered to allow a larger grouping of numbers. The results point out similar trends to those already discussed for Figure 55.

A more complex picture of the levels of phosphorus in Alberta soils can be seen in Figure 57 where a smaller circle has been used to produce the final results. Figure 57 shows the distribution of phosphorus in Alberta on all crops using farm samples that have been analyzed since 1962. Only 3 of the 8 concentric circles composing the mapping function were used with the weightings of 3, 2, 1. The final result is a fairly difficult map to discern rapidly. This type of map, however, is quite useful for delineating specific areas of high or low values. Local trends that are often masked by the larger, regional mapping function can be more readily distinguished on such a detailed map.

The area summaries have pointed out that although phosphorus is more closely associated to soil areas, it does vary from year to year. A comparison of phosphorus maps for each year should point out this variation if it does exist. Figures 58, 59, 60, 61 are maps of Alberta showing the distribution of phosphorus on fallow, stubble and grass-legume crops for the sampling years 1968-69, 1967-68, 1966-67 and 1965-66, respectively. A comparison of these maps shows that differences do exist from year to year. The sampling year 1968-69 appears to contain the lowest levels of phosphorus, while that of 1965-66 appears to have a high phosphorus content. Similar results have been noted in the area summaries. Generally, most regions tend to vary similarly from year to year. This same trend did not appear in the nitrate maps. Areas of equal levels (as defined by the contour lines), however, tend to take on a different shape from year to year. For





Figure 57. Map of Alberta showing distribution of phosphorus on all crops, using a smaller circle, for the time period: Apr., 1962-Mar., 1969. Circle weights: 3 2 1 0 0 0 0 0.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

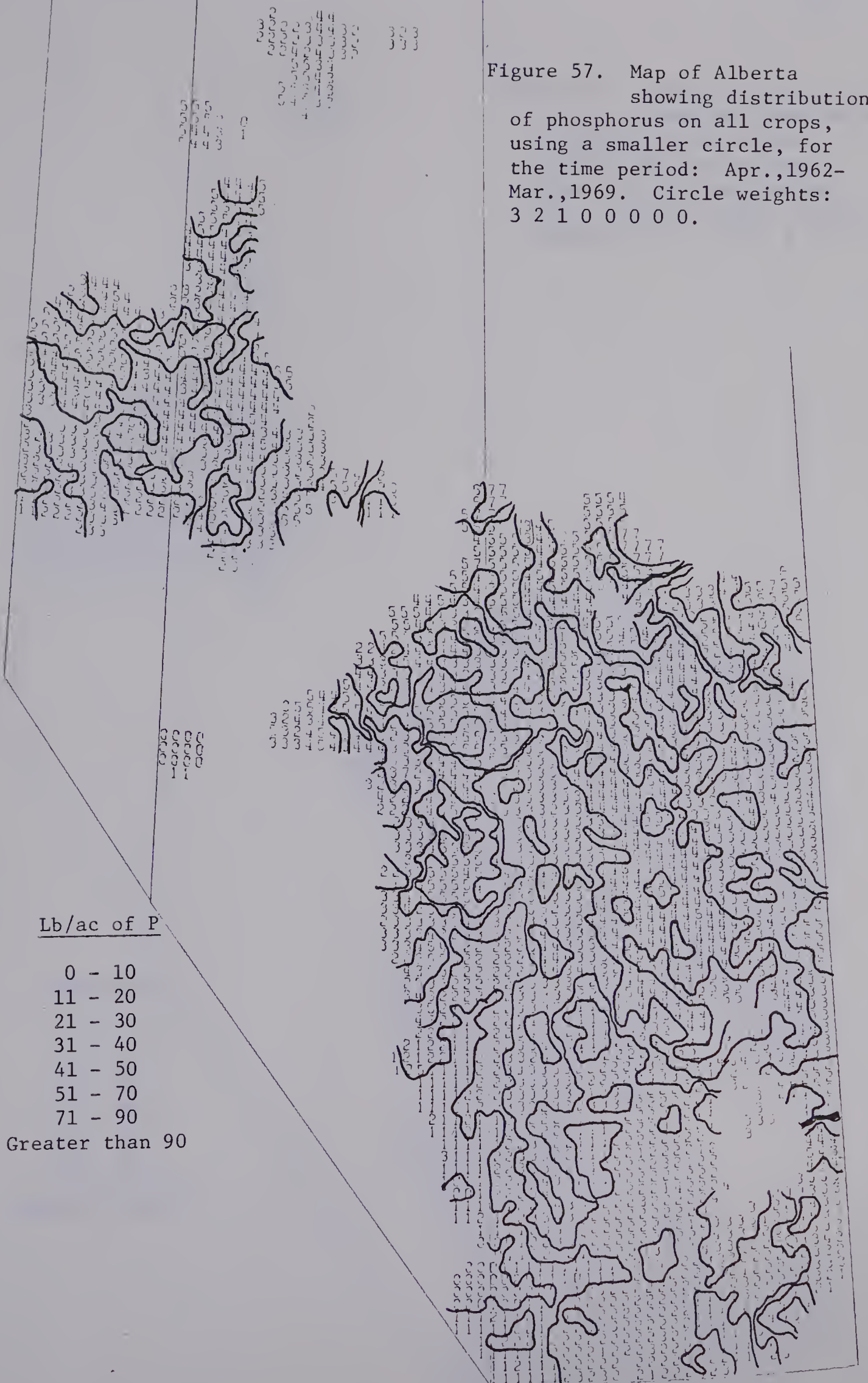




Figure 58. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1968-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

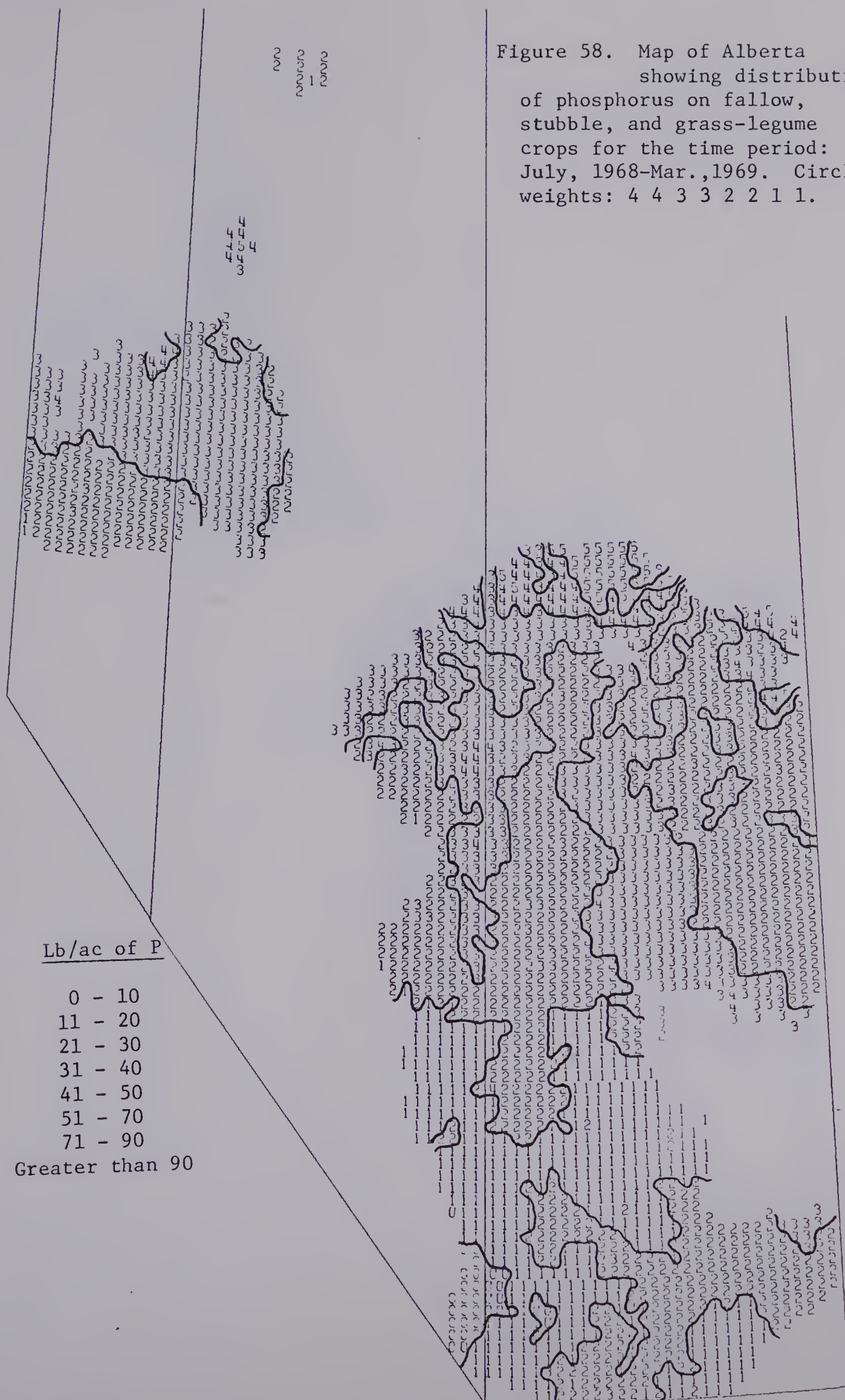






Figure 59. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1967-Mar., 1968. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

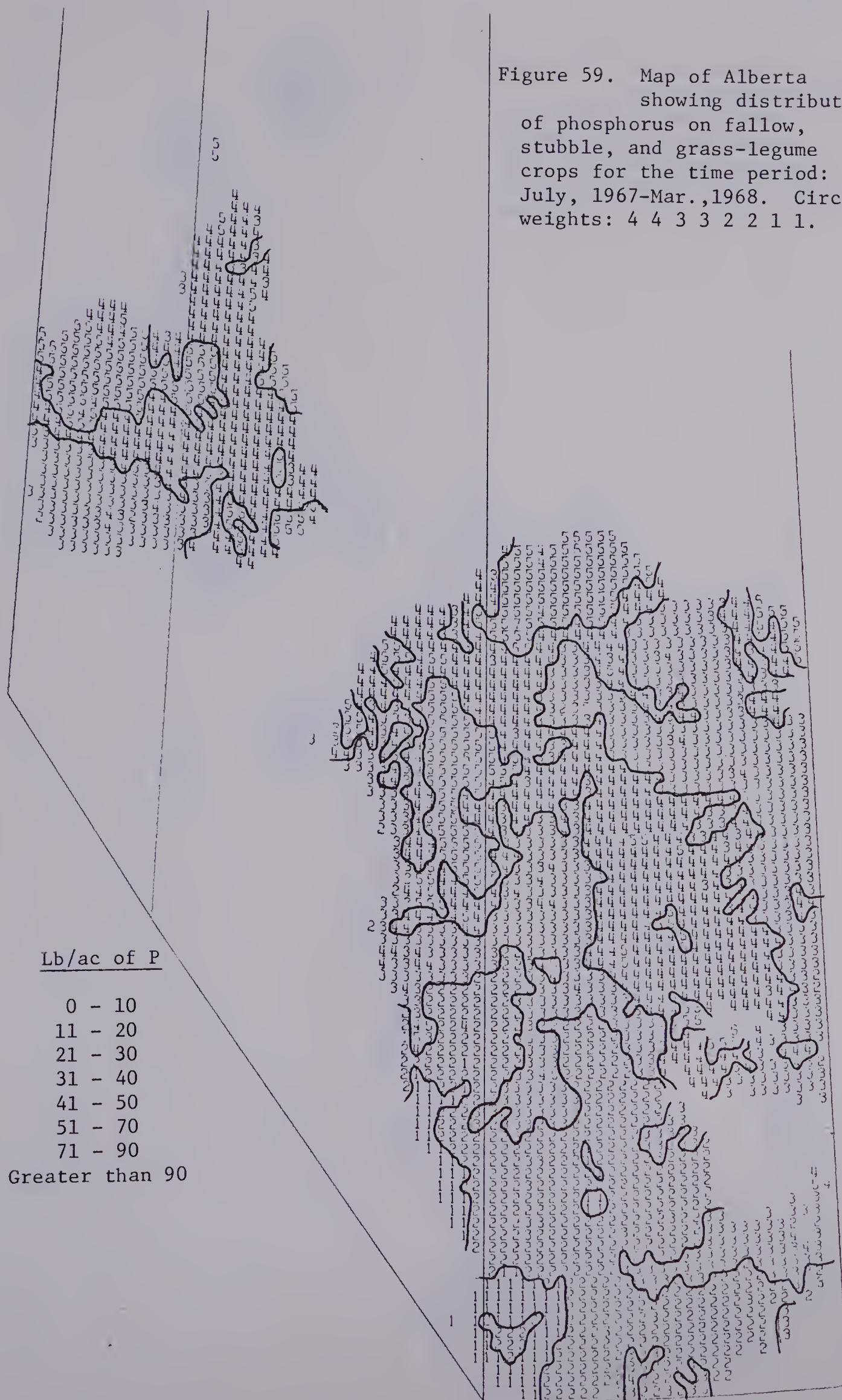




Figure 60. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1966-June, 1967. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90

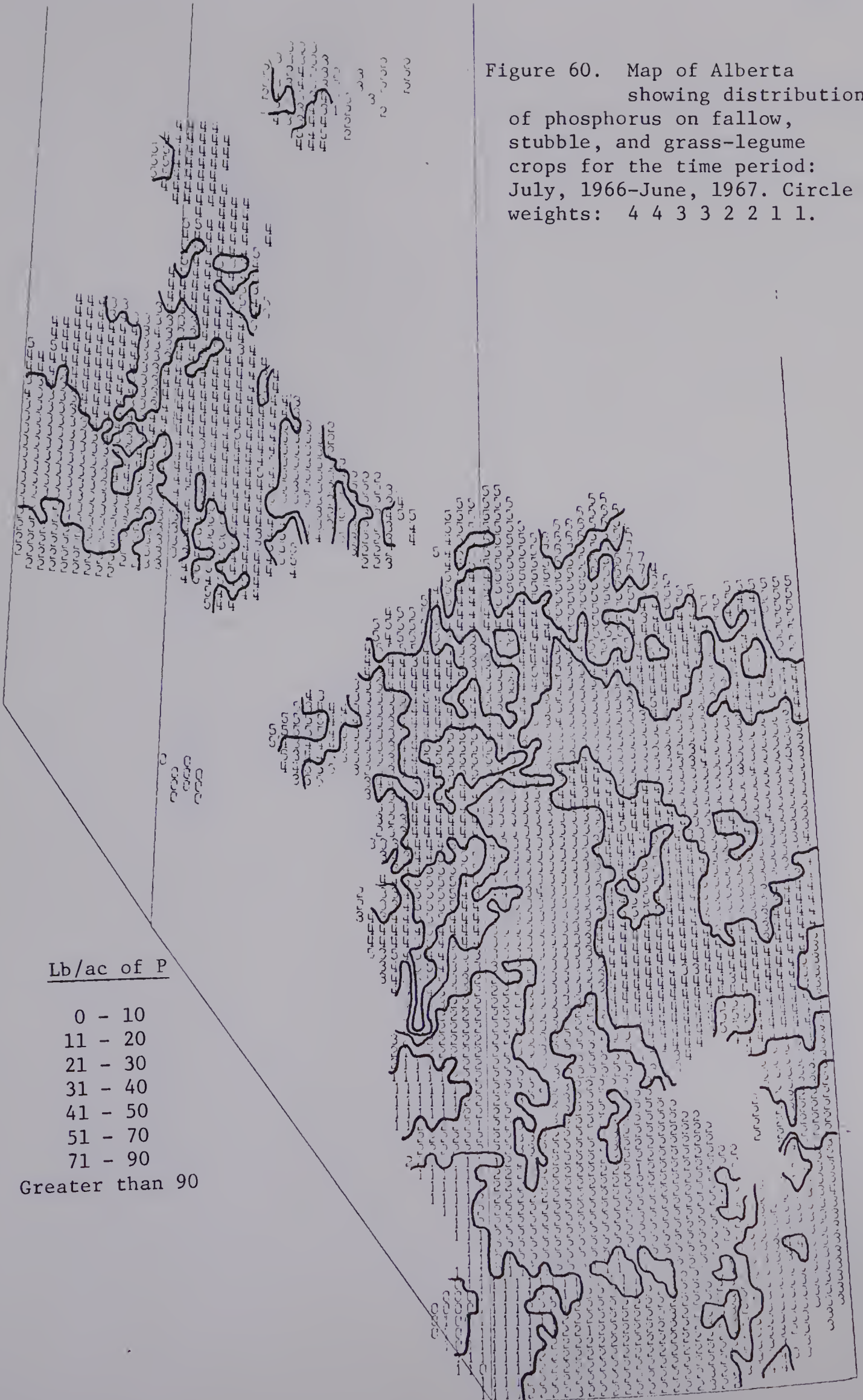
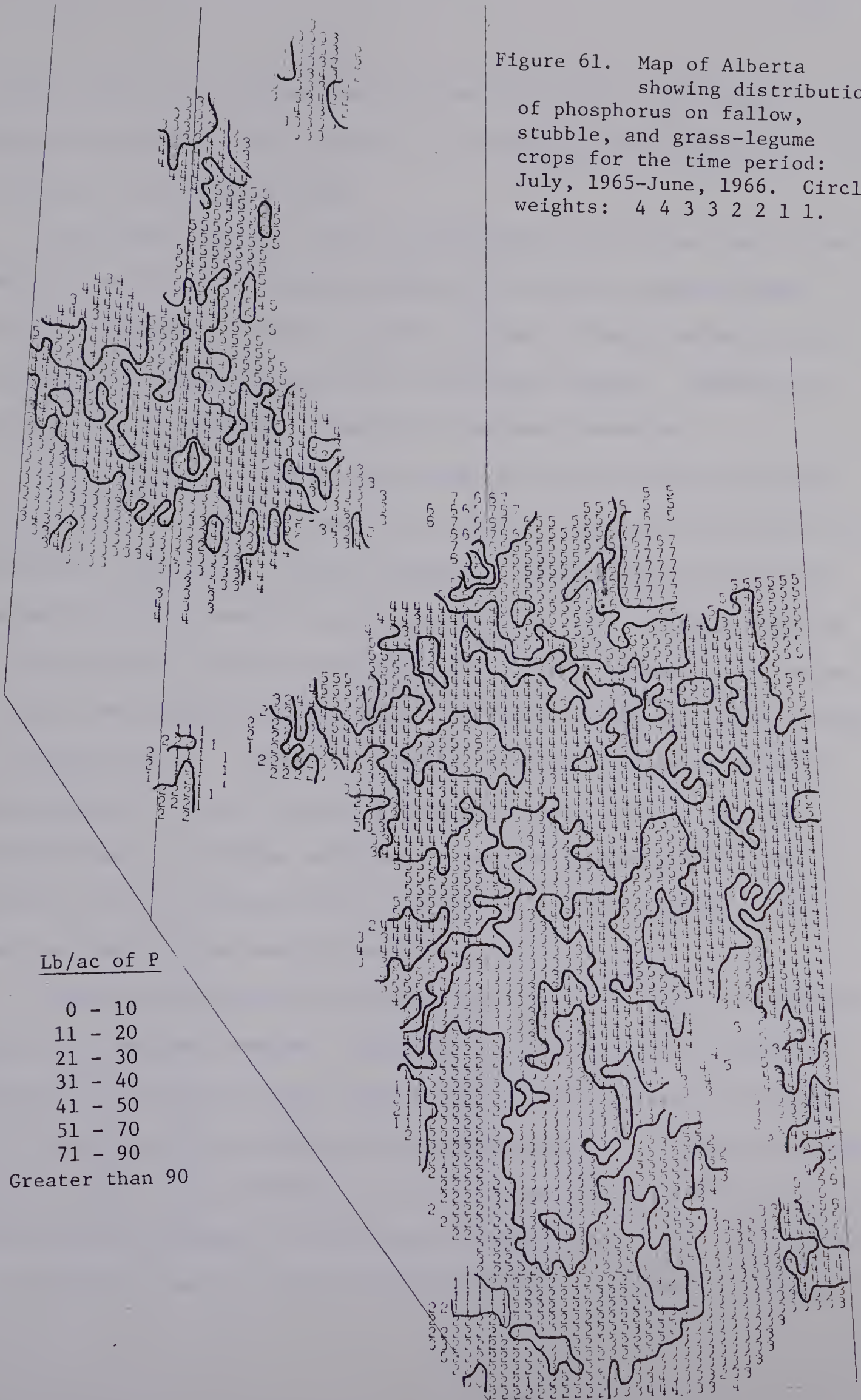






Figure 61. Map of Alberta showing distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period: July, 1965-June, 1966. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of P
0	0 - 10
1	11 - 20
2	21 - 30
3	31 - 40
4	41 - 50
5	51 - 70
6	71 - 90
7	Greater than 90







example, the area to the southeast of Edmonton takes on a different shape and sometimes level each year. A comparison of the maps will illustrate the different shapes.

The results show that levels of phosphorus do vary from year to year and that usually this variation affects all areas in a similar manner. Levels of phosphorus, however, appear to be more closely related to soil types or soil-climatic areas than to any yearly changes. Identical results have been previously presented in the area summaries.

(d) Potassium - - The coded average values of potassium for each township where at least two samples are present have been plotted for the province in Figure 62. The map includes all samples that have been received since April, 1965, when the new potassium test was introduced, up to March, 1969. The crop grouping was decided upon from previous results for the area summaries showing little difference between potassium content on different crops. The map illustrated in Figure 62 shows some fairly distinguishable trends considering that the values displayed are not mapped values. A definite decrease in potassium content can be seen as one goes from the southern part of the province to the northern. Such a map can almost be contoured in its present form.

However the application of the mapping function to the potassium data makes this task much simpler. Figure 63 is a map of Alberta showing the potassium distribution on all crops for all the years since April, 1965. The final contoured map shows an interesting property - - potassium content appears to be closely related to the soil-climatic areas in Alberta. The Brown soil zone appears to have levels of around 801 to 1000 lb/ac. There is a distinct area projecting through the middle of the Brown soil zone



Figure 62. Map of Alberta showing potassium values on all crops for townships with 2 or more samples available. Time period: Apr., 1965 - Mar., 1969.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000

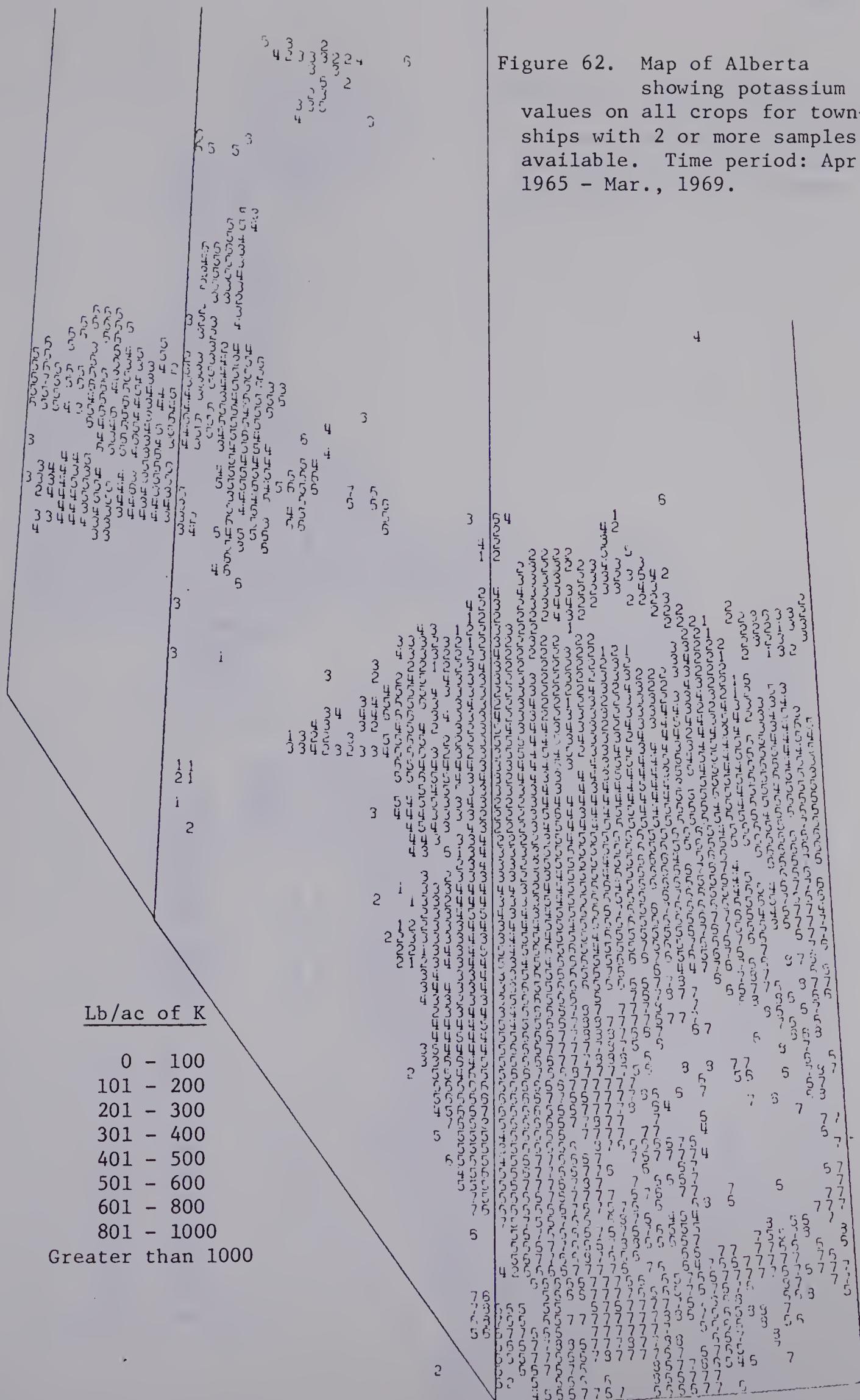
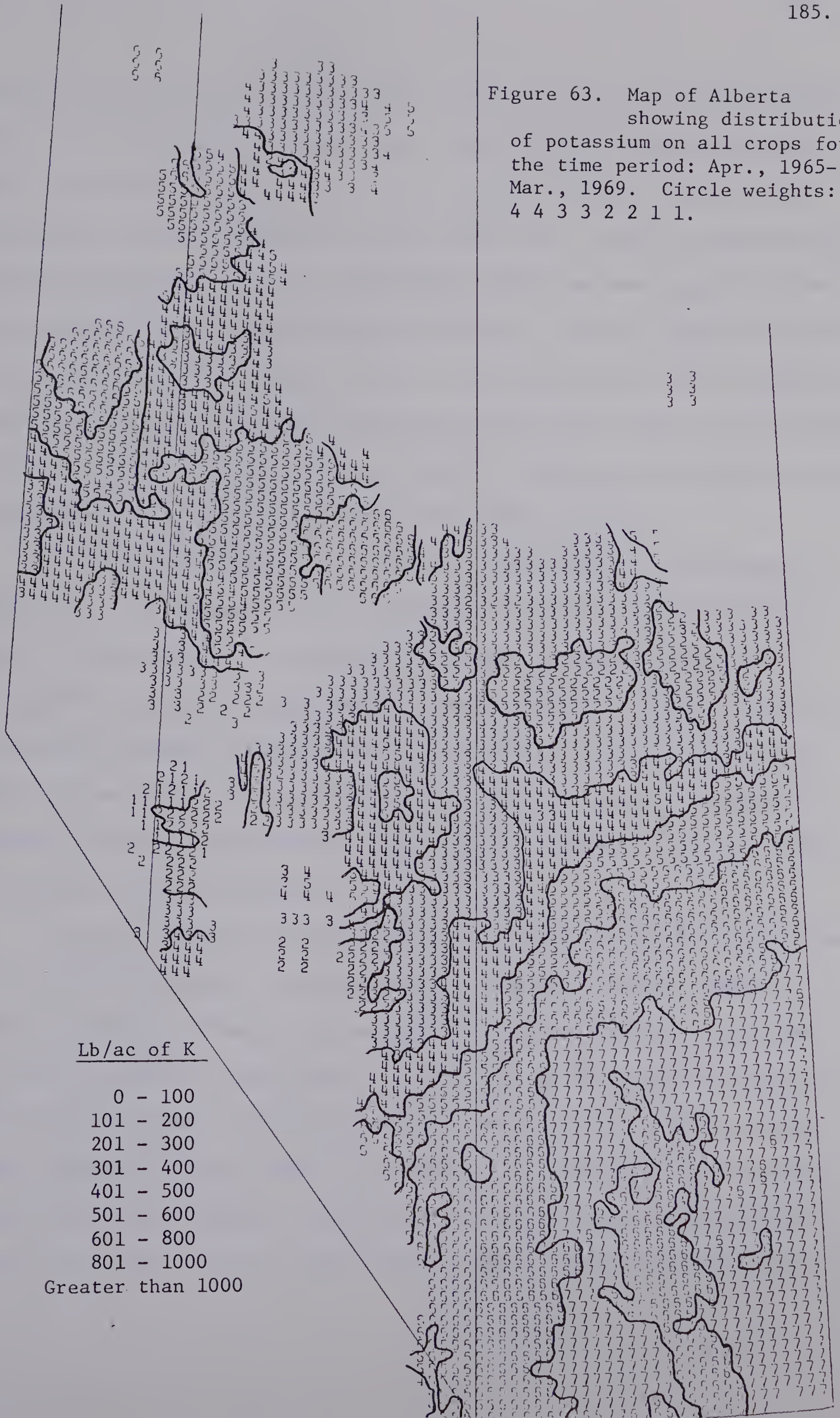






Figure 63. Map of Alberta showing distribution of potassium on all crops for the time period: Apr., 1965-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000





showing a decrease in potassium content with averages between 601 and 800 lb/ac. The Dark Brown soil zone shows values similar to this particular area. The Thin Black soil zone shows averages of 501-600 lb/ac and the Black soil zone shows averages of 401 to 500 lb/ac. The Gray Wooded soils usually show averages of 301 to 400 lb/ac. There are some pockets to the north and northeast of Edmonton where fairly low values of potassium occur with averages between 201 and 300 lb/ac. The Peace River region generally shows levels similar to those previously noted for the Black and Thin Black soil zones, however pockets that are similar to the levels recorded for the Dark Brown zone as well as the Gray Wooded soils do appear.

A more detailed map of the potassium distribution is illustrated in Figure 64. All the data available since April, 1965 has been used to produce the final outcome. In addition, a smaller mapping function with 3 concentric circles with weightings 3, 2 and 1 was used. Similar trends as previously discussed can generally be pointed out. Low potassium areas, such as those to the north and northeast of Edmonton are more discretely depicted, but the relationship of the potassium levels to the soil zones is no longer as easy to visualize.

The area summaries and the work of other researchers have pointed out that potassium does not vary substantially from year to year. However, a comparison of the various years would be of interest in this study, even if only to verify the above point. Figures 65, 66, 67, and 68 are maps of Alberta showing the potassium distribution on all crops for the sampling years 1968-69, 1967-68, 1966-67, and 1965-66. A comparison of these maps shows that usually similar areas with similar levels of potassium exist from year to year. No two years, however, have exactly similar contours.





Figure 64. Map of Alberta  
showing distribution  
of potassium on all crops,  
using a smaller circle, for  
the time period: Apr., 1965-  
Mar., 1969. Circle weights:  
3 2 1 0 0 0 0 0.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000

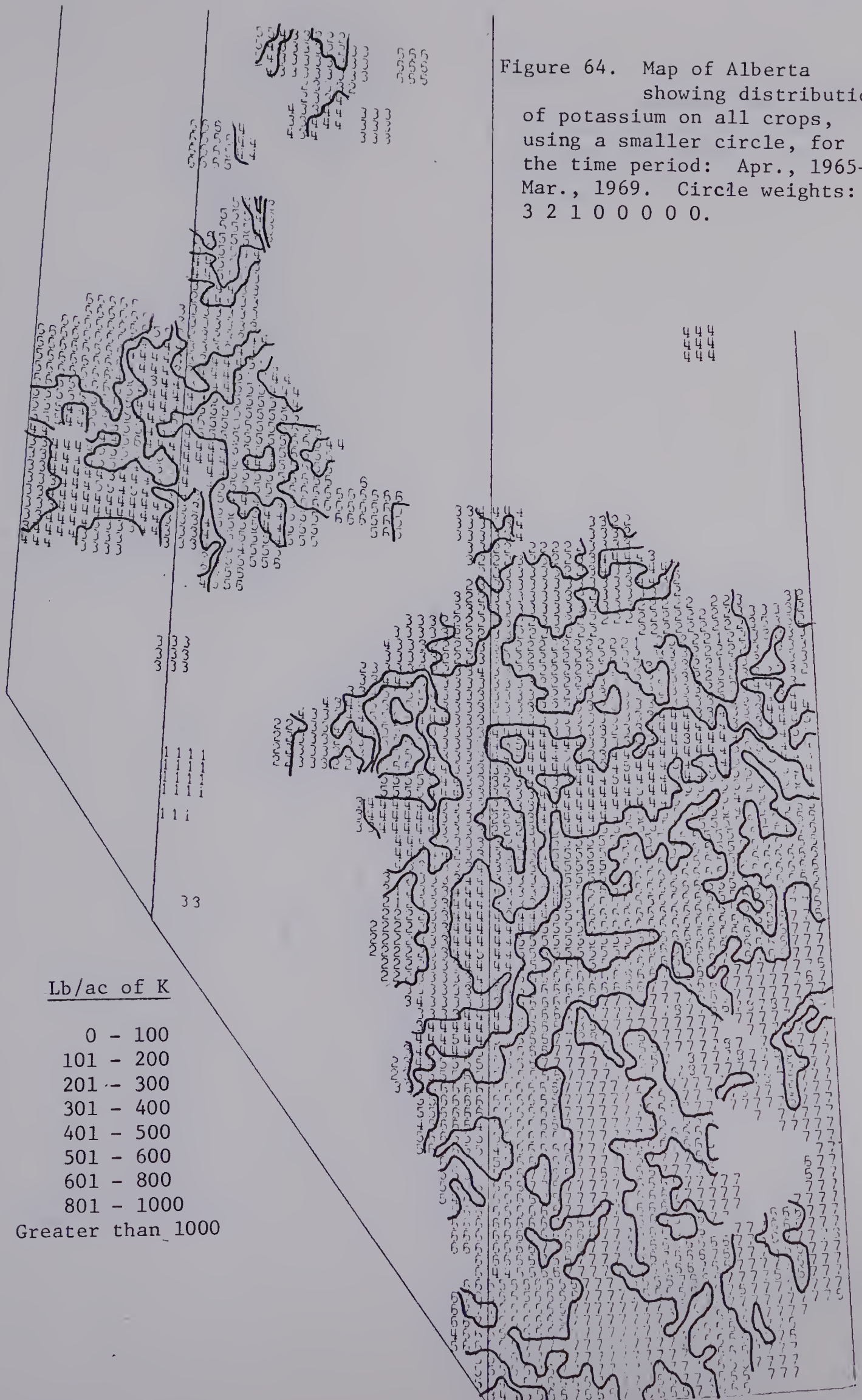




Figure 65. Map of Alberta showing distribution of potassium on all crops for the time period: July, 1968-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000

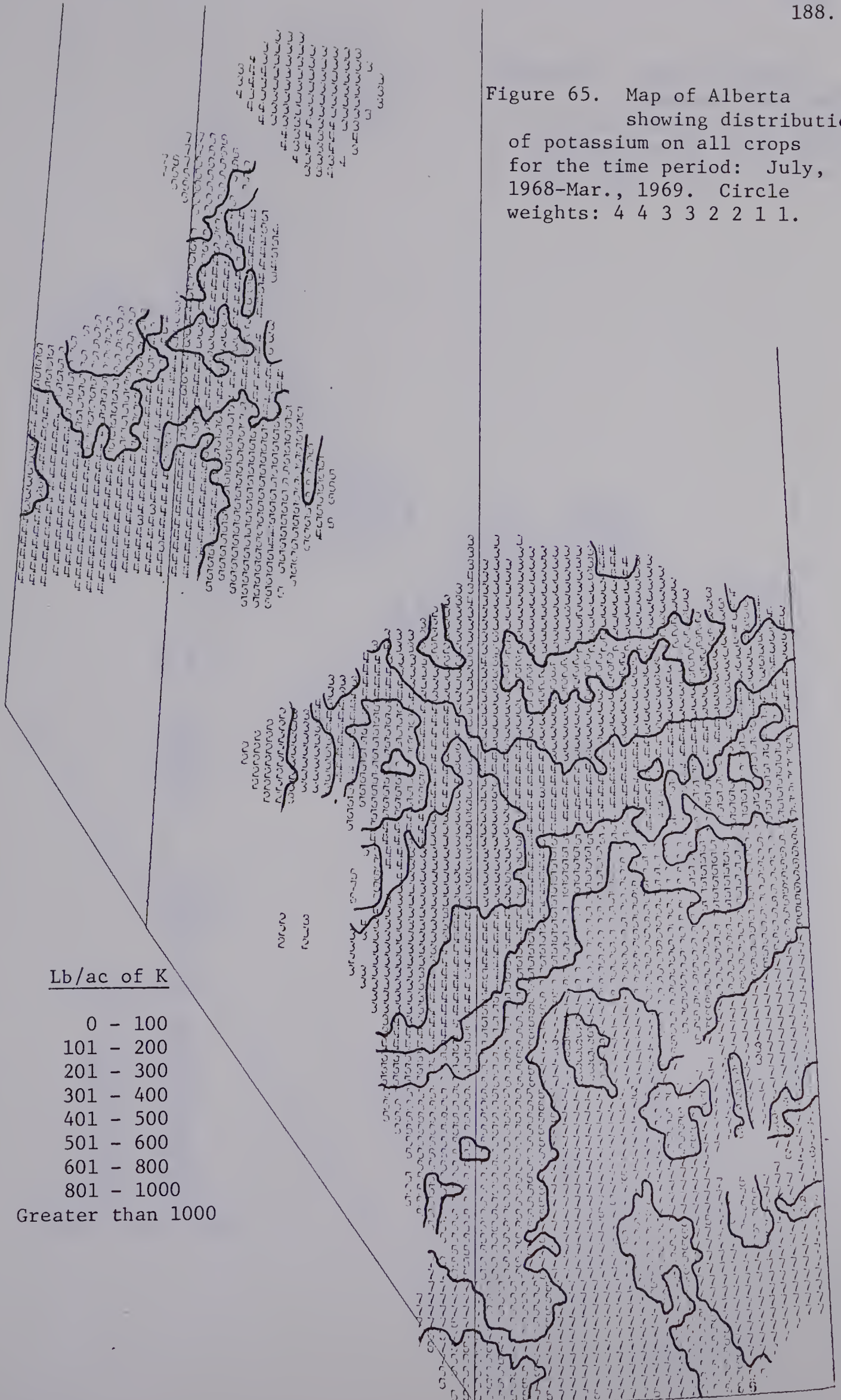






Figure 66. Map of Alberta  
showing distribution  
of potassium on all crops for  
the time period: July, 1967 -  
June, 1968. Circle weights:  
4 4 3 3 2 2 1 1.

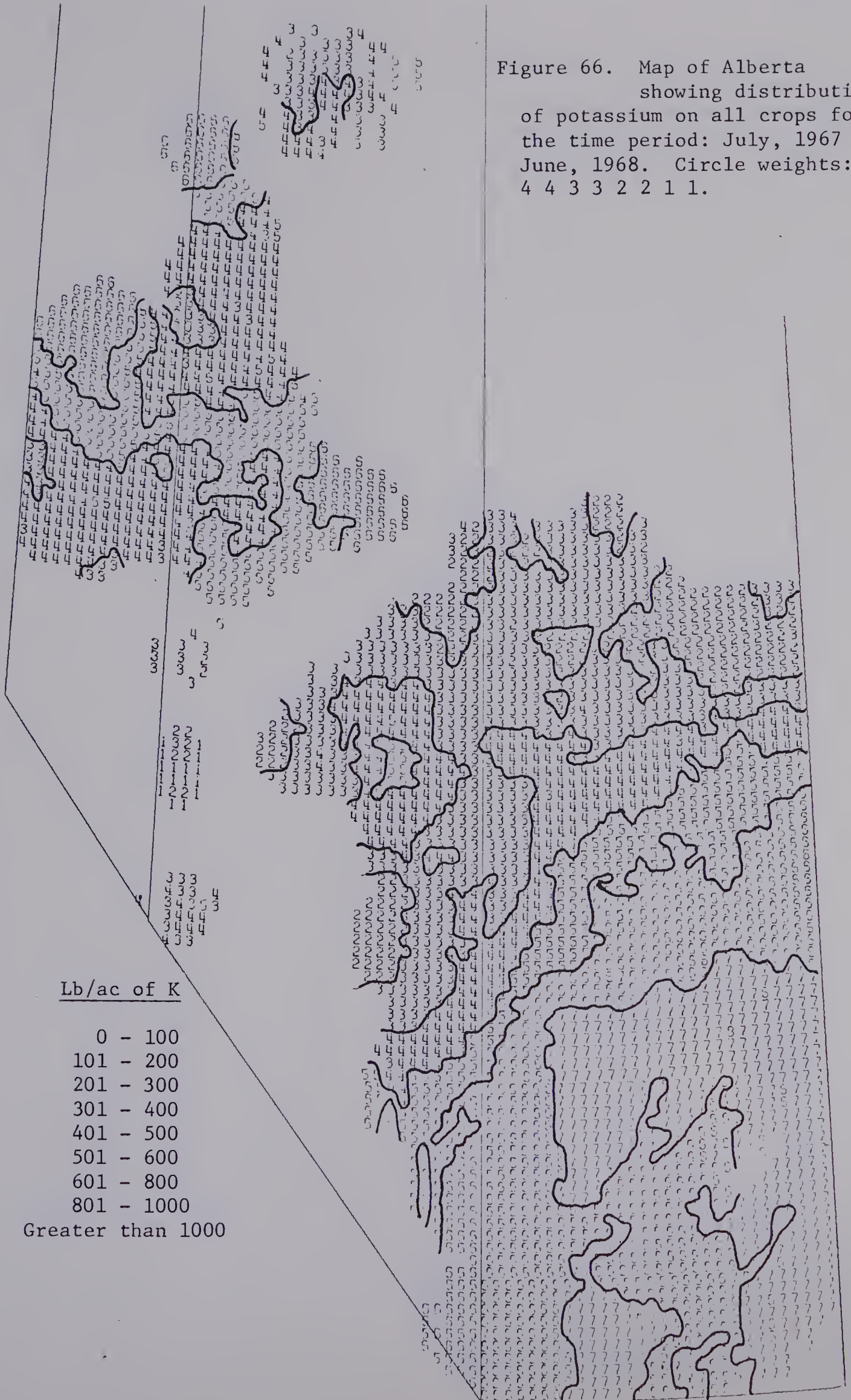






Figure 67. Map of Alberta showing distribution of potassium on all crops for the time period: July, 1966 - June, 1967. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000

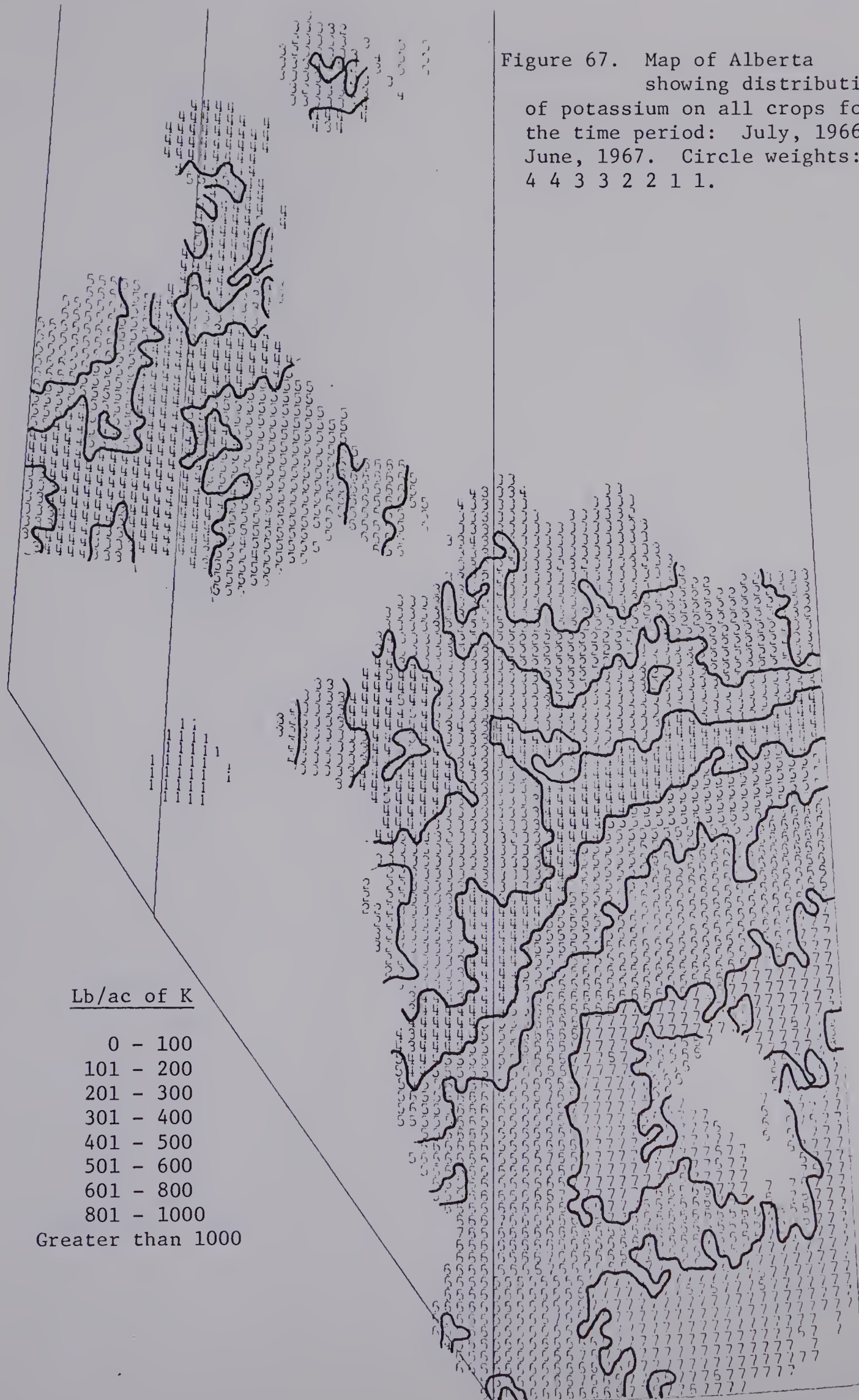
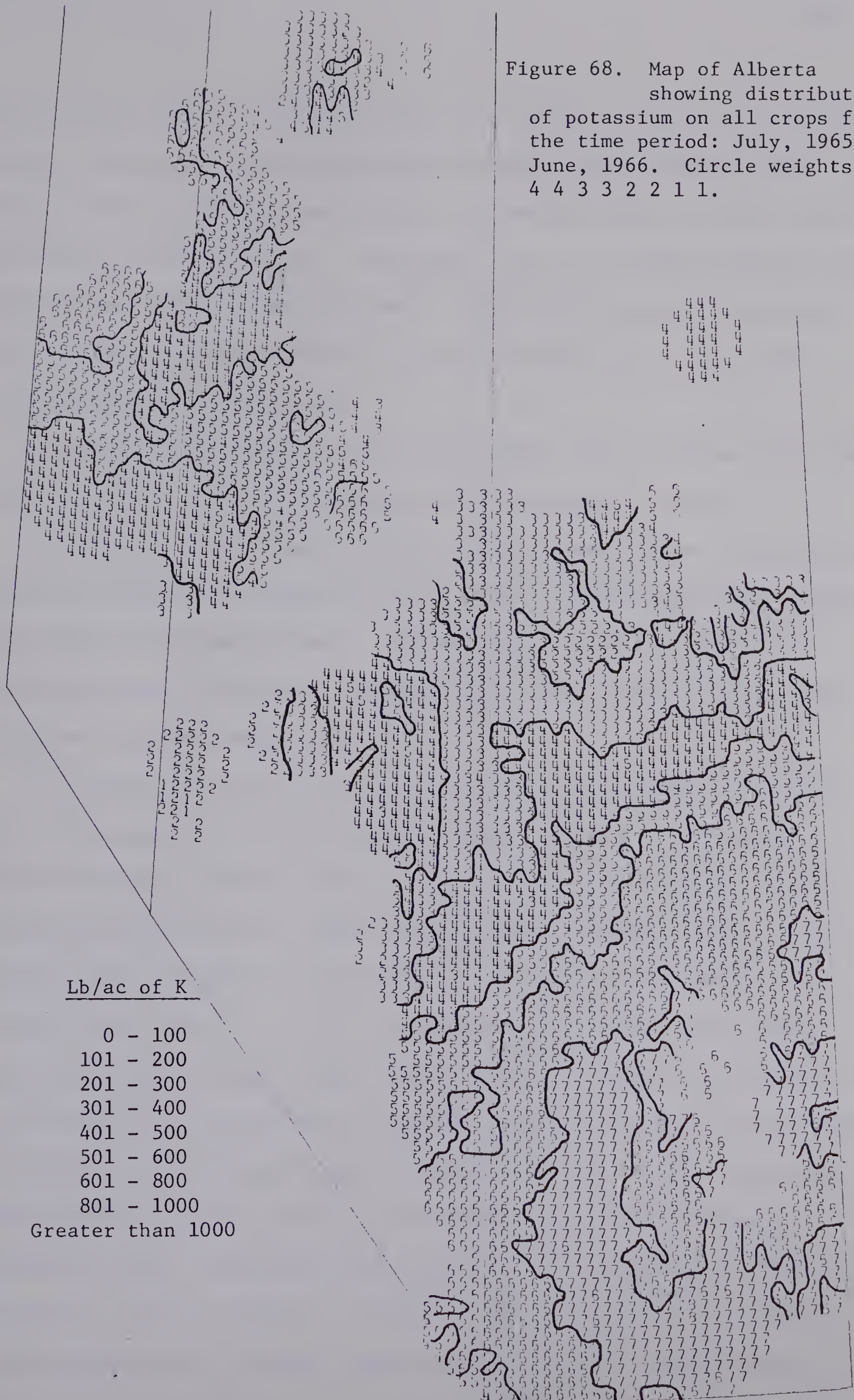




Figure 68. Map of Alberta showing distribution of potassium on all crops for the time period: July, 1965 - June, 1966. Circle weights: 4 4 3 3 2 2 1 1.

Code	Lb/ac of K
0	0 - 100
1	101 - 200
2	201 - 300
3	301 - 400
4	401 - 500
5	501 - 600
6	601 - 800
7	801 - 1000
8	Greater than 1000







Certain areas tend to fluctuate in size and shape from year to year. For example, the area of declining potassium values south of Bassano and usually between Lethbridge and Medicine Hat tends to take on quite a variable shape and size each year. Similarly, the low potassium areas to the south and northeast of Edmonton seem to change shape and size each year. Part of this alteration from year to year may well be due to the lower sampling density in these areas.

Despite the small fluctuations in potassium levels for some areas from year to year, no general trend along this line appears to exist.

(e) Soil pH - - A map of Alberta showing coded numbers representing specific pH ranges in townships where at least 2 samples have been received since 1962 is shown in Figure 69. Figure 69 is useful for picking out observed averages for specific townships, but not as useful for depicting trends on a regional basis.

The mapping of pH data, as has been carried out in Figure 70, gives a very clear regional picture of the pH trends. Figure 70 is a map of Alberta displaying relative levels of pH for the A.S.F.T.L. data for all crops and years from April, 1962 to March, 1969. The map shows that the southern areas of Lethbridge and Medicine Hat have the more basic soils with pH values from 7.1 to 7.5. The Calgary and Red Deer areas show a slight drop with averages ranging from 6.6 to 7.0. There is an island of soils directly to the southeast of Camrose with very low pH values averaging from 5.6 to 6.0. This same area corresponds to the area where relatively high phosphorus values occur, the area that approximately outlines the Solonchic soils. The region south of Edmonton and Lloydminster shows a pH between 6.1 and 6.5, usually. The area to the north, however, tends to show an increase with averages ranging from 6.6 to 7.0. The Peace River



Figure 69. Map of Alberta showing pH values on all crops for townships with 2 or more samples available. Time period: Apr., 1962-Mar., 1969.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5

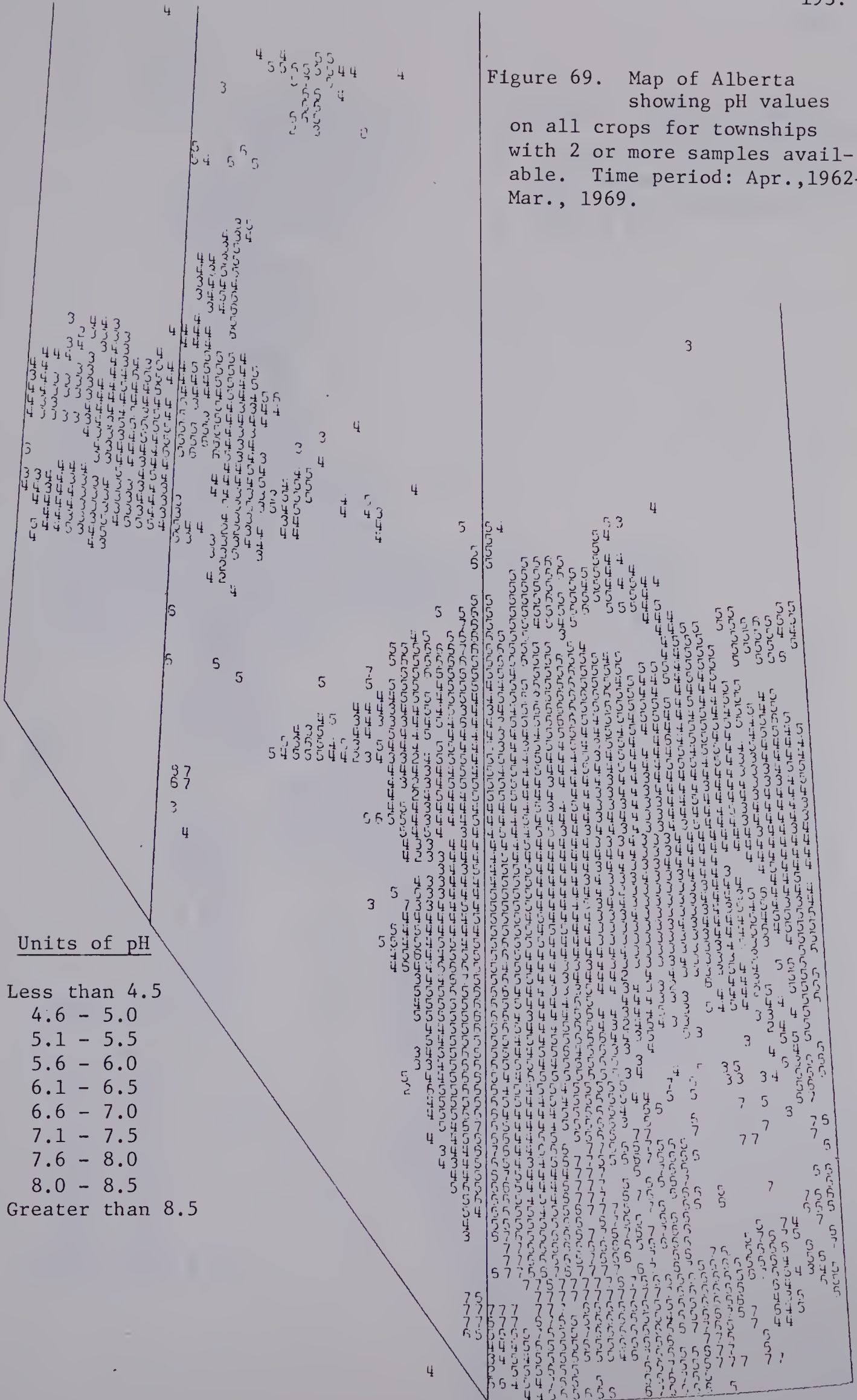
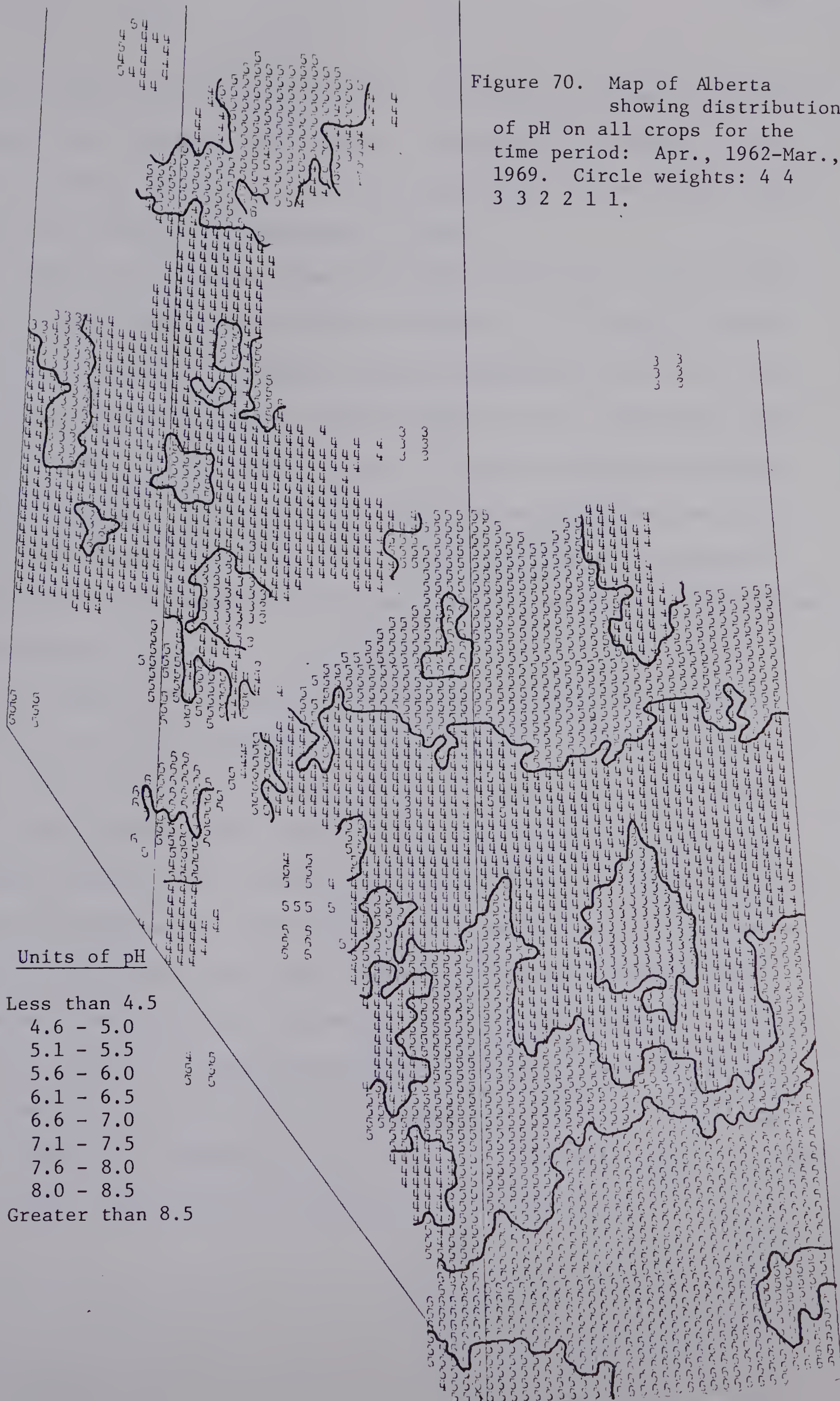






Figure 70. Map of Alberta showing distribution of pH on all crops for the time period: Apr., 1962-Mar., 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5





region displays pH values of around 6.1 to 6.5 with a decrease at Grande Prairie and to the northwest of Grande Prairie, and an increase in the Fort Vermilion country. It is interesting to note that the Cypress Hills area can be readily distinguished on the map.

Figure 71 is a map of Alberta showing the distribution of pH values for the same crops and time period as in Figure 70, but using a smaller circle for mapping. The mapping function involved the use of 3 of the 8 available concentric circles with weightings: 3, 2, 1. The final result is a more detailed map displaying more of a spotty pattern but pointing out local areas of high and low pH values that were previously masked by the larger mapping function. Maps of this nature are quite valuable in depicting areas of local variation and possibly areas where future studies are warranted

A comparison of pH values over the years is displayed in Figures 72, 73, 74, and 75 for the sampling years 1968-69, 1967-68, 1966-67, and 1965-66. The results show that very little difference exists in pH from year to year for any particular locality. However, the shape and size of the general areas do seem to differ from year to year. This same trend has been previously noted in the potassium and phosphorus maps. For example, the area of low pH values to the southeast of Camrose shows definite changes in shape and size from one year to the next. Similarly, the area to the northwest of Grande Prairie tends to fluctuate from year to year. It is interesting to note that in the sampling year 1968-69 a fairly large area north and west of Lethbridge appears to have an unusually high pH of around 7.6 to 8.0. This trend is not apparent in the other years.





Figure 71. Map of Alberta showing distribution of pH on all crops, using a smaller circle, for the time period: Apr., 1962-Mar., 1969. Circle weights: 3 2 1 0 0 0 0 0.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5

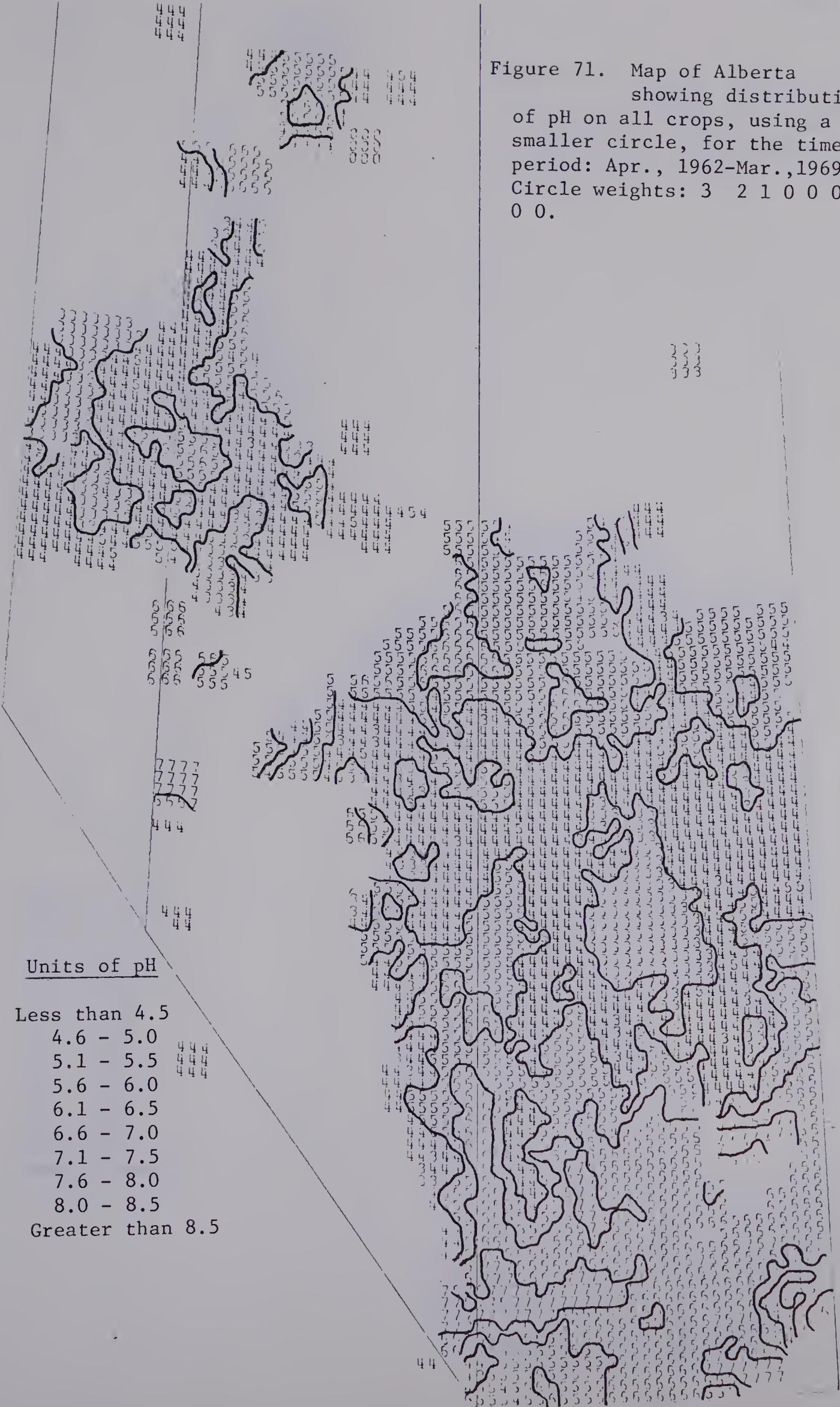






Figure 72. Map of Alberta showing distribution of pH on all crops for the time period: July, 1968-June, 1969. Circle weights: 4 4 3 3 2 2 1 1.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5

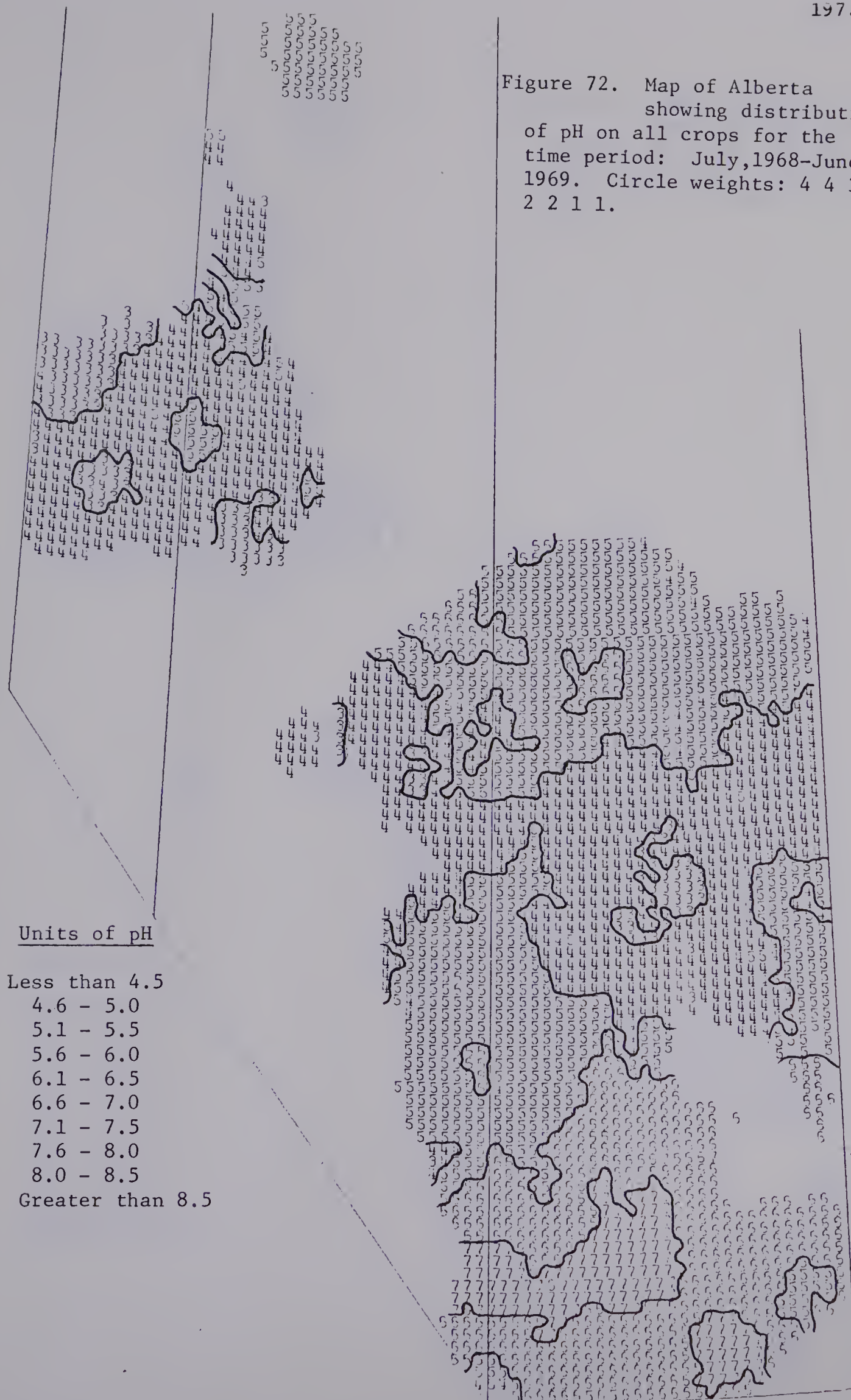
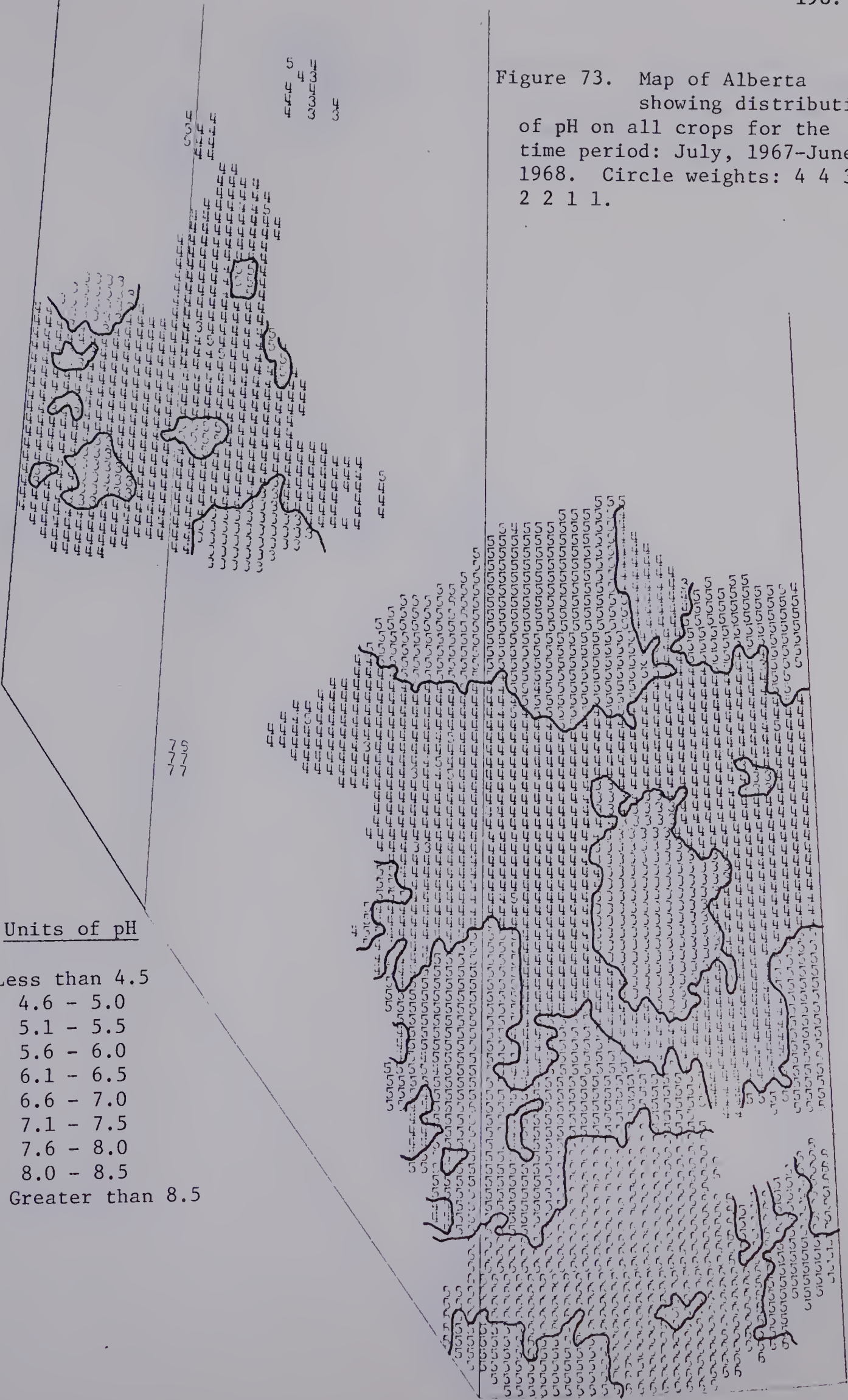




Figure 73. Map of Alberta showing distribution of pH on all crops for the time period: July, 1967-June, 1968. Circle weights: 4 4 3 3 2 2 1 1.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5







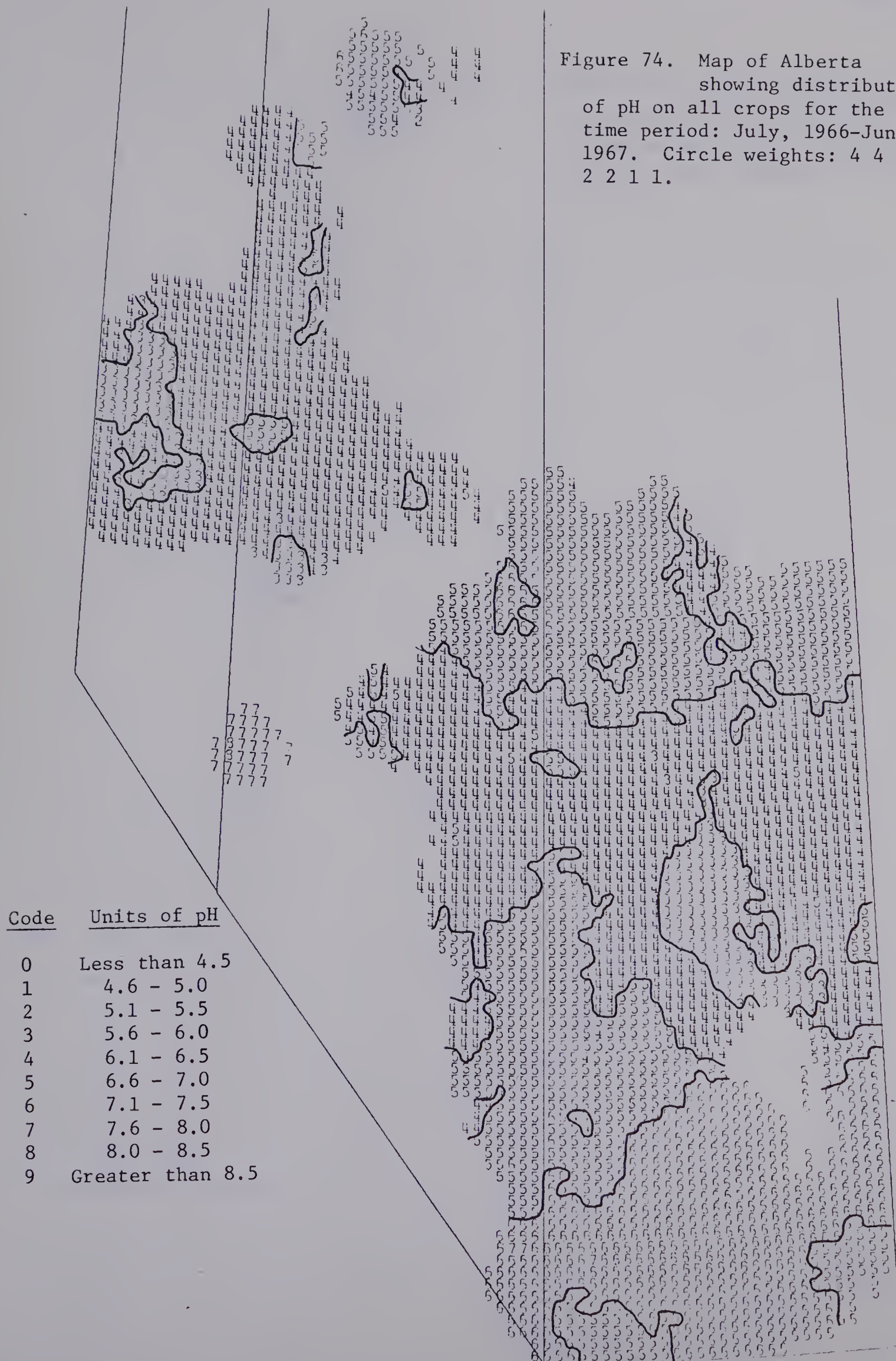
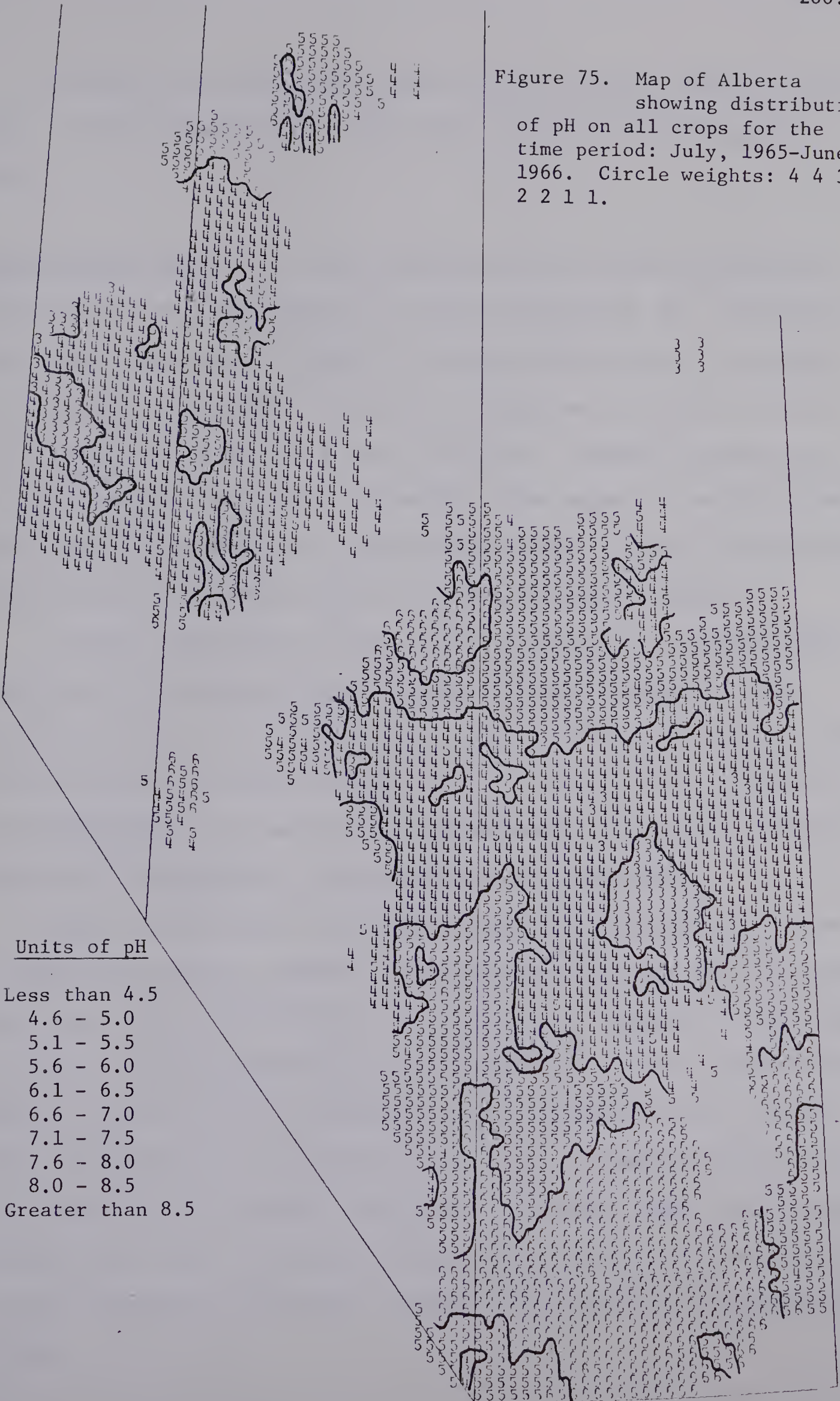




Figure 75. Map of Alberta showing distribution of pH on all crops for the time period: July, 1965-June, 1966. Circle weights: 4 4 3 3 2 2 1 1.

Code	Units of pH
0	Less than 4.5
1	4.6 - 5.0
2	5.1 - 5.5
3	5.6 - 6.0
4	6.1 - 6.5
5	6.6 - 7.0
6	7.1 - 7.5
7	7.6 - 8.0
8	8.0 - 8.5
9	Greater than 8.5







In summary, then, the results show that pH is closely related to soil areas. Annual variation does occur, but no apparent extreme variations occur.

County nutrient maps: The circular function used to map the province was applied to the area surrounding the County of Camrose No. 22. The section rather than the township was used as the basic mapping unit. The purpose of this study was two-fold. Firstly, it allowed one to obtain an idea of the variation that exists on a more local basis. Secondly, it served as an evaluation of the possibility of mapping other areas in a similar manner. County maps were produced for phosphorus, potassium, and pH. No map was produced for nitrate nitrogen, because of inadequate sample numbers.

In each of the three maps that are presented on the following pages, exactly the same mapping control variables were used, except for time periods and crops. All 8 circles of the circular mapping function were used with the following assigned weights: 4, 4, 3, 3, 2, 2, 1, 1. The size of the mapping function was such that it automatically encompassed all samples within an 8 mile radius for each final map value calculation.

Figure 76 is a map of the County of Camrose showing the distribution of phosphorus on fallow, stubble, and grass-legume crops for the time period from April, 1962 to March, 1969. The value for each isopleth is illustrated on the map. The phosphorus map is quite variable, but trends can be depicted. There is a general trend for the east side of the County to show a higher level of phosphorus than the adjacent west side. A comparison with Figures 55, 56, and 57, the provincial maps, shows very similar trends, but with considerably less local variations. The high levels noted on the east side of the map correspond to the general location of the Solonchic soils discussed earlier.



Figure 76. Map of the County of Camrose showing isopleths of phosphorus content on fallow, stubble, and legume-grass crops. Time period: Apr., 1962-Mar., 1969.

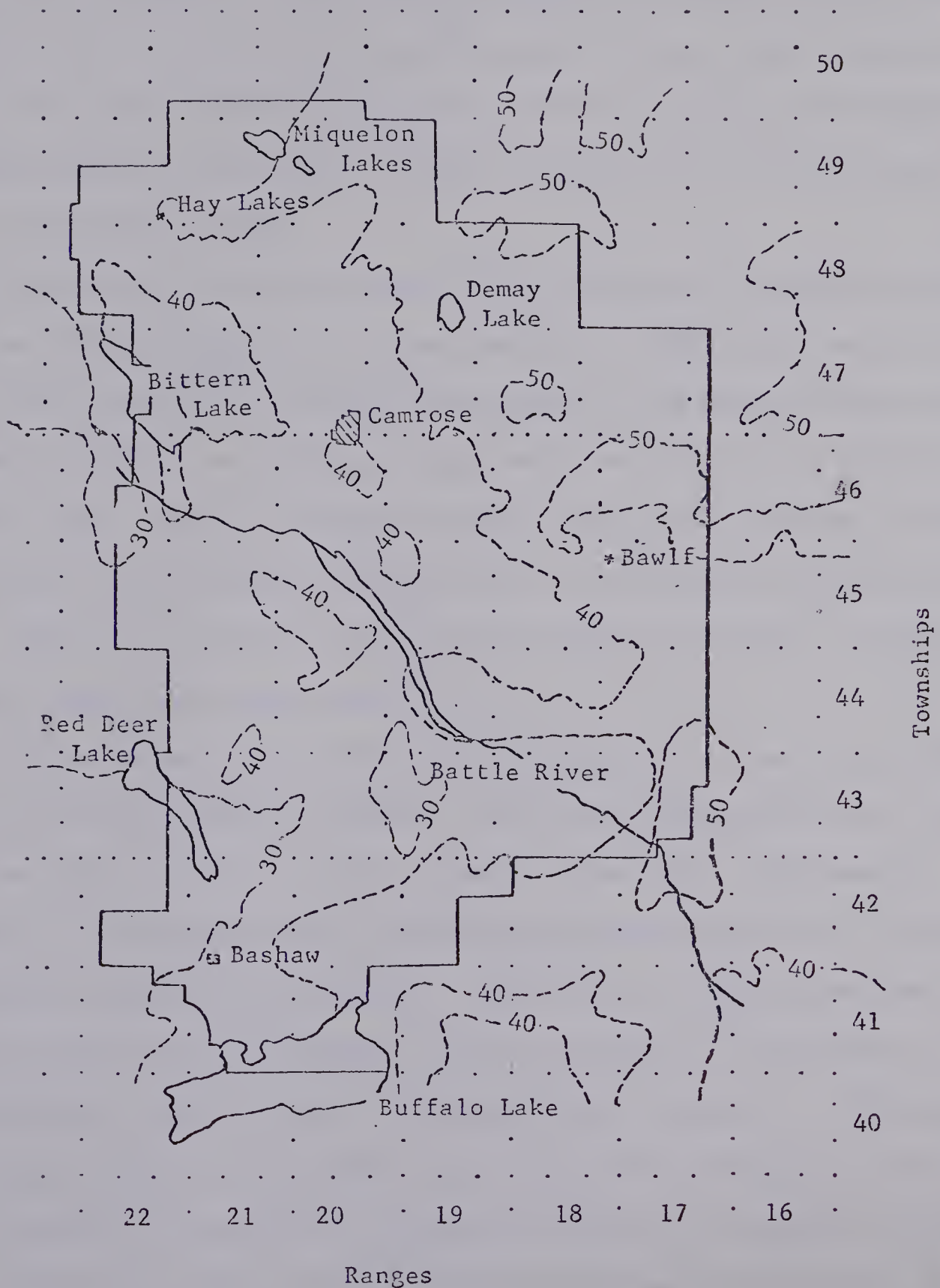




Figure 77 is a map of the County of Camrose showing the distribution of potassium on all crops for all the A.S.F.T.L. samples received from April, 1965 up to March, 1969. The resulting map shows definite trends. There is an increase in the potassium content as one goes from the north to south end of the map. The southern part of the map shows averages around 601 to 700 lb/ac, while the area to the north of Camrose shows averages ranging from 401 to 500 lb/ac. Figures 63 and 64, the provincial maps, show very similar trends.

The pH map presented in Figure 78 for the County of Camrose includes all the soil test data on all crops from 1962 to 1969. The results indicate very little variation in pH with the majority of the County having averages between 5.6 and 6.5 . Because of this small variation, it is difficult to depict trends. However, the map does show that the more acidic pH values are found to the northeast of Camrose and in the southeast corner of the map. Figure 71, a pH map of the province obtained by using a small mapping circle, shows very similar trends.

The county maps have a definite advantage over the large provincial maps in that they can be utilized to locate more accurately areas of high and low levels of soil nutrients. However, they have a disadvantage, working on a sectional rather than a township basis, and that is the lack of adequate samples in many places. Where too few samples exist for a reliable map value to be calculated, contour lines must be approximated. This has been done to some extent in the maps already presented. The shortage of samples, however, will improve with time as more samples are sent in to the laboratory and analyzed. The results indicate that mapping soil nutrient levels for a county, using a section as the basic mapping unit, is quite feasible in areas of fairly high sampling density.





Figure 77. Map of the County of Camrose showing isopleths of potassium content on all crops. Time period: Apr., 1965-Mar., 1969.

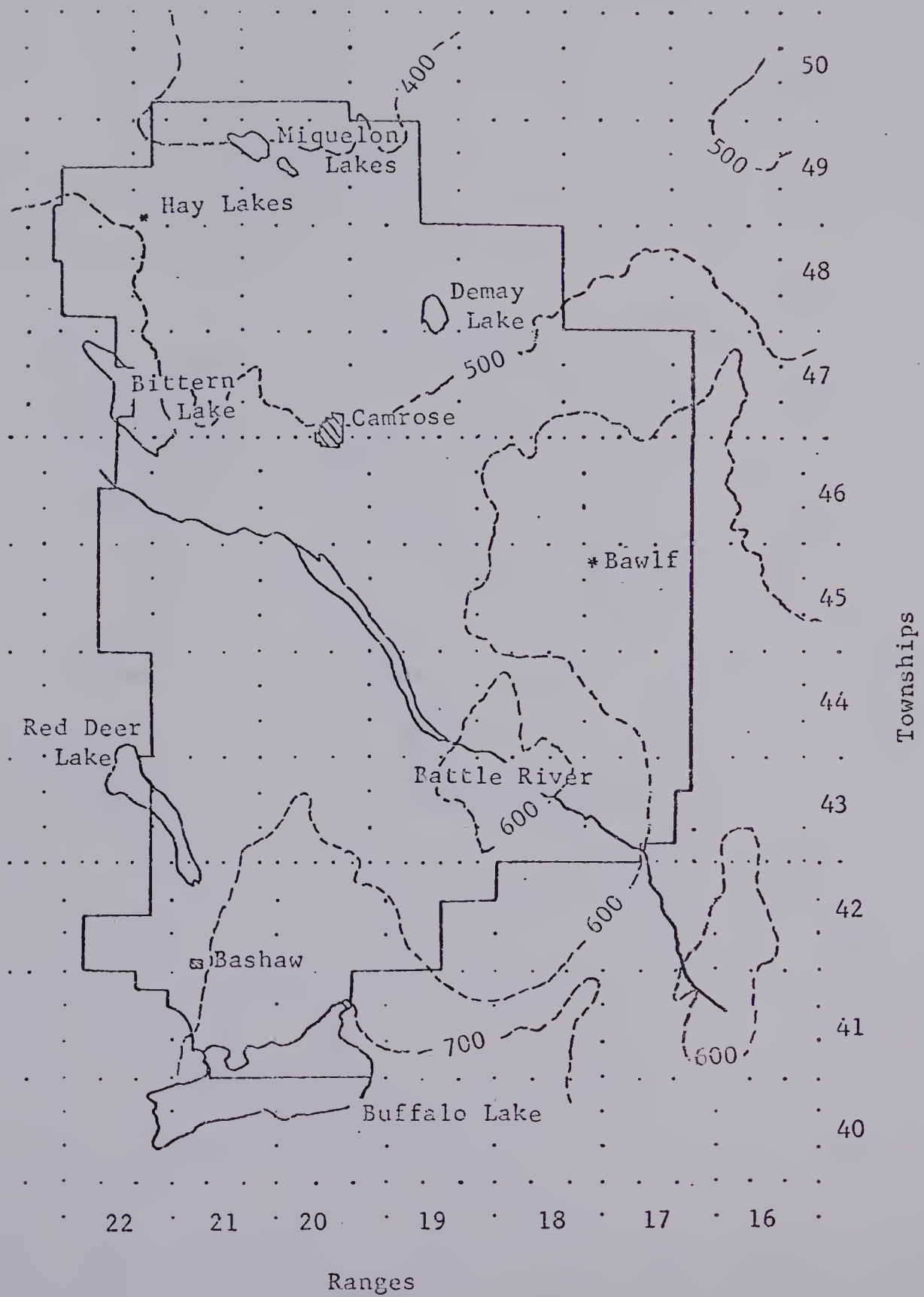
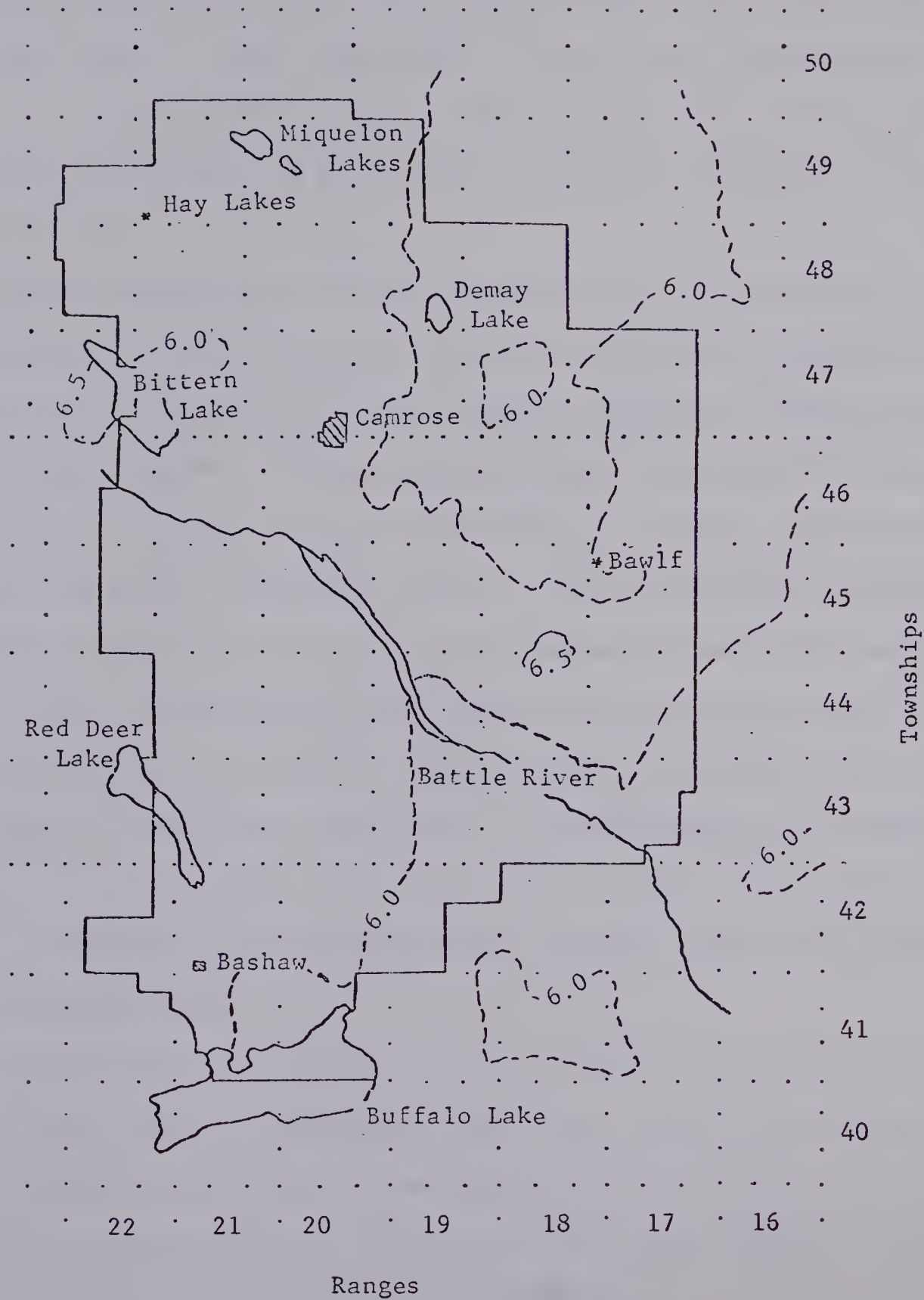




Figure 78. Map of the County of Camrose showing isopleths of soil pH on all crops. Time period: Apr., 1962-Mar., 1969.





Field mapping: The Cameron farm data previously used to estimate the sampling accuracy of various nutrients, was sampled in a grid-like pattern, thus, enabling the rapid application of the circular mapping function. The maps of the province show large regional trends. The county nutrient maps show trends on a more local basis, but still over a sizeable area. The Cameron farm data allows one to observe trends within a field, where the regional component will most likely be strongly influenced by slopes and soil types.

The basic mapping unit for the field was a cell, 46 x 40 feet. A total of 6 concentric circles with equal weighting composed the mapping function.

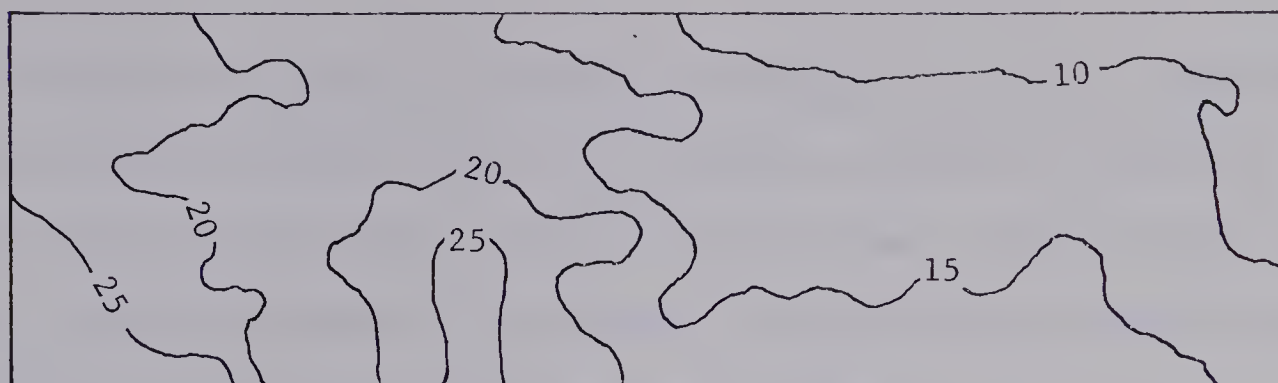
The mapped results for nitrate nitrogen, phosphorus, potassium and pH are shown in Figure 79. Nitrate nitrogen shows a decrease of 15 lb/ac, from 25 lb/ac in the southwest corner of the field, to 10 lb/ac in the northeast corner of the field. The slight rise to 25 lb/ac in the more central part of the map covers part of the area of the eastward slope of the main hill in the field. The phosphorus and pH maps display nicely the peak of this hill. The phosphorus content is variable, even in the east half of the field where no significant slopes occur. Potassium shows an increase in the east half of the field. Soil pH tends to be very stable. No relationships to soil type can be inferred since a detailed soil classification of the field was not made.

The maps point out the large amount of variation that can exist in a field. Trends within an individual field can be easily differentiated using the mapping function, but the incorporation of fertilizer practices to fulfill the deficiencies as indicated by the trends would not only be difficult, but impractical with present technology.

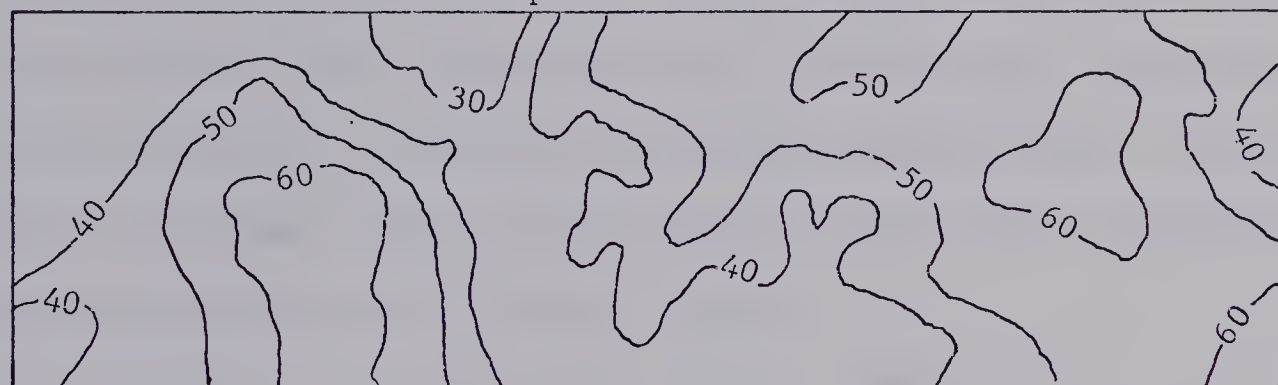


Figure 79. Nutrient maps of the Cameron farm field showing distribution of nitrate nitrogen, phosphorus, potassium, and pH.

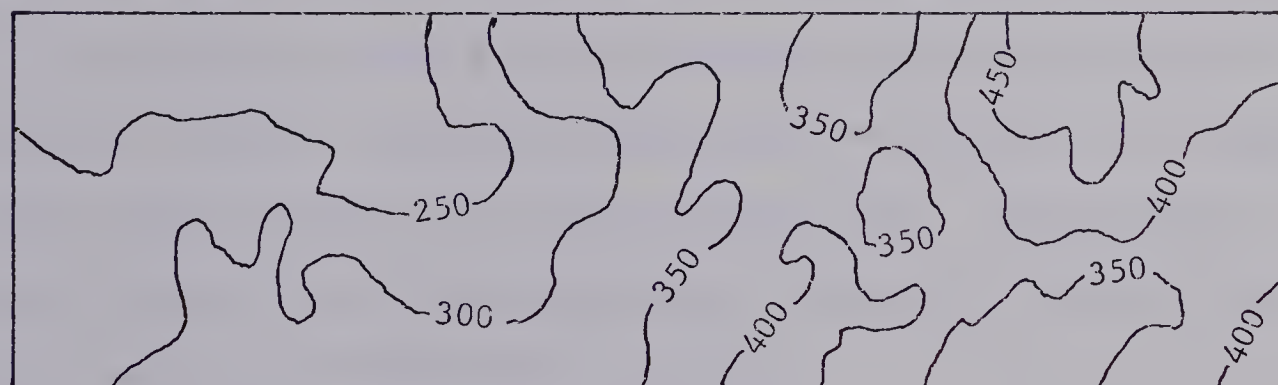
## Nitrate nitrogen



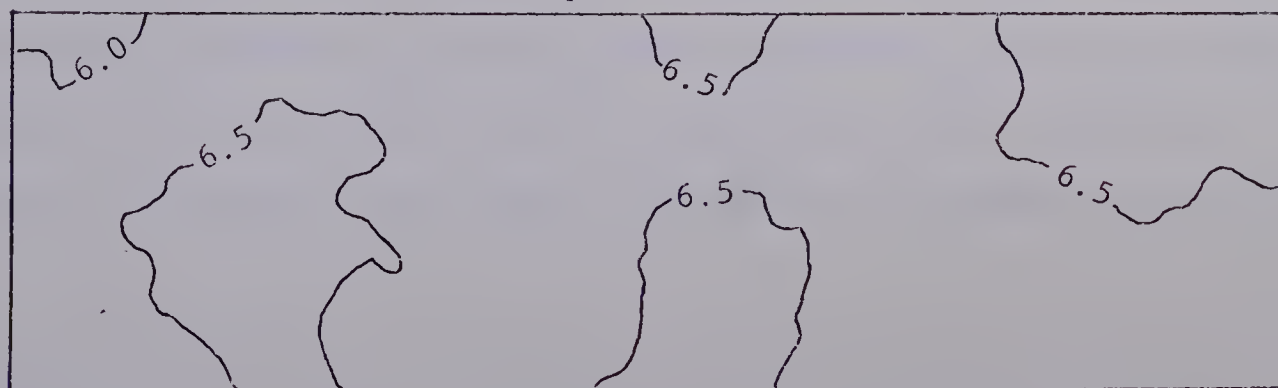
## Phosphorus



## Potassium



## Soil pH







### Results of Prediction of Unknowns

Through use of simulation techniques "probable" values were predicted for nitrogen and potassium results that had been previously recorded up to, but not beyond certain limits, called "ceiling values". Without this prediction, the final results for any summary or map would have shown extreme bias, moreso in the case of potassium than nitrogen.

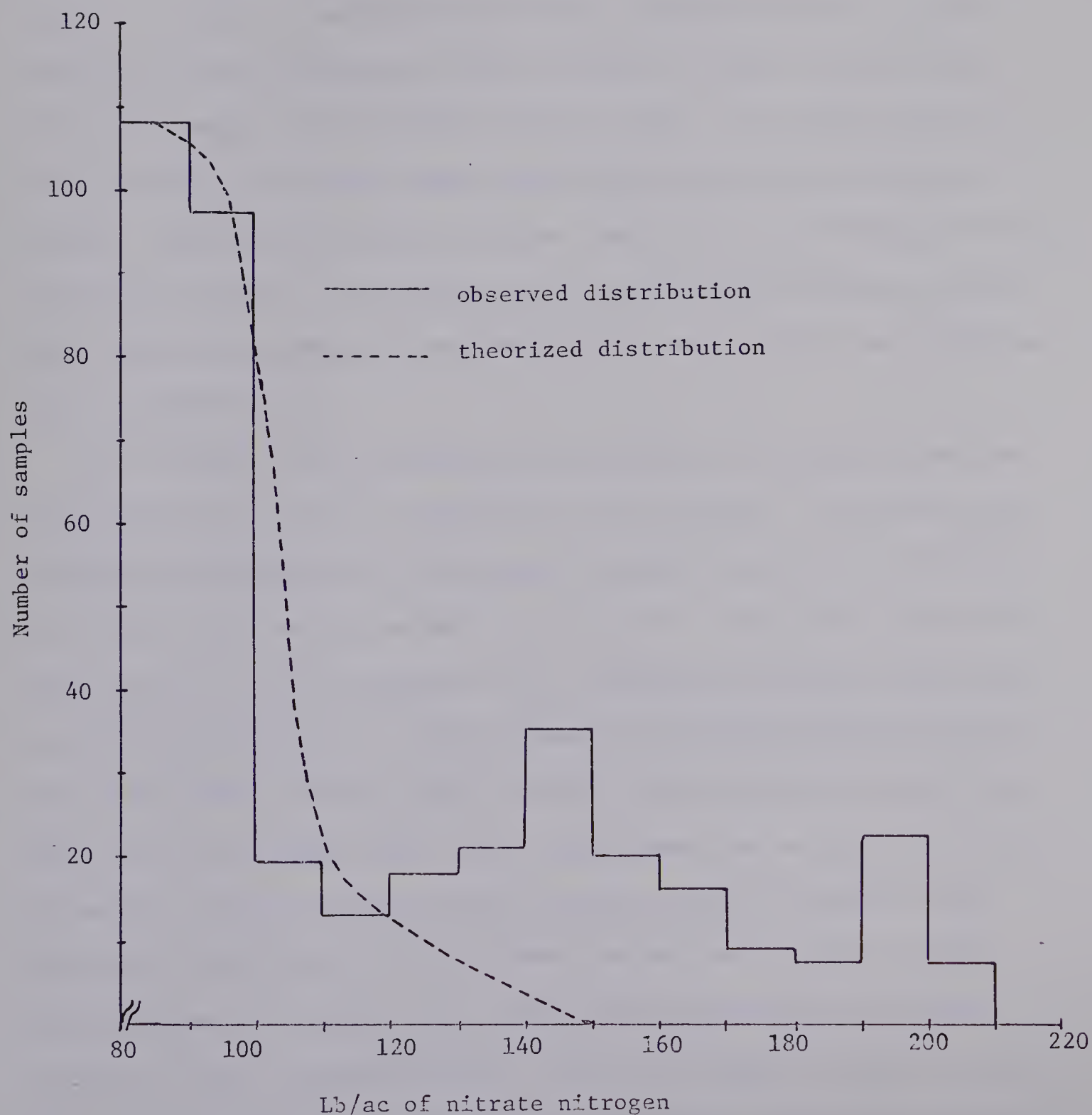
The results for nitrate nitrogen show, for the samples received before August, 1968, that approximately 2.5% or 1400 samples were recorded as 80 + with Area 6 showing 3.2% or 500 from a total of 15,830 samples recorded as 80 +. The partialness that might be incurred by "ceiling values" would be small due to the large number of other samples present, but a reasonable prediction would produce more accurate results.

The results of the analyses of the 387 soil samples which were previously recorded as 80 + are shown in Figure 80. The solid-lined histogram outlines the number of samples within each 10 lb/ac interval from 80 to 210 lb/ac. The results show that 54% of the samples have values below 100 lb/ac and the remainder are spread unevenly over an interval of 110 lb/ac with minor peaks occurring at 145 lb/ac and 195 lb/ac. The unevenness associated with the tail of the distribution could not be linked to the soil area, texture, or organic matter content.

Extremely high values of nitrate nitrogen representing field means for the province are quite rare and most likely unrealistic. The occurrences of such anomalies are usually linked to improper collection, drying, or storage of the soil sample prior to acquisition by the Laboratory. Therefore, an arbitrary upper limit of 150 lb/ac was decided upon for the



Figure 80. Observed and theorized distributions of high nitrate nitrogen values previously recorded as 80 +. Total number of samples used: 387.





development of the model distribution. Any value beyond this point would usually be considered as erroneous.

A best estimate of the type of distribution that would most likely be formed, had no erroneous values appeared, is represented on the graph in Figure 80 by the dashed line. This model distribution was used to simulate a value for unknown nitrate levels recorded as 80 +. If this distribution was simulated an infinite amount of times to derive known values for samples with recorded "ceiling values", the results would be such that 76% of simulated values would fall between 81 and 100 lb/ac, inclusive; 18% between 101 and 120 lb/ac, inclusive; and 6% between 121 and 150 lb/ac, inclusive. No results for the final values predicted will be shown, since they are not as important as the other results and should be quite straightforward.

The "ceiling values" employed for high levels of potassium are discussed under the heading "Materials" in this thesis. An analysis was conducted to determine the percentage of samples by soil area for which any of the three "ceiling values" were recorded. The results, shown in Table XX, point out the importance of a reasonable prediction being made for these unknown values. Table XX shows the percentage distribution for each of the three "ceiling values" by soil area using all the data from April, 1965 up to March, 1969. The results show that usually over 50% of the samples from the southern areas have upper limits placed on them. The use of these results for any summary or map would create a large amount of bias. The other areas show smaller percentages of recorded "ceiling values", but nevertheless, substantial enough to cause some bias. Table XX illustrates the importance of predicting a reasonable result for





Table XX. Percentage distribution of potassium "ceiling values" by soil area for the A.S.F.T.L. data collected in time period of April, 1965 to March, 1969.

Area	* <u>"Ceiling values"</u>			<u>Total percent with "ceiling values"</u>	<u>Total no. of samples</u>
	<u>601</u>	<u>801</u>	<u>999</u>		
1	38	10	7	55	1938
2	50	7	6	63	3059
3	47	7	6	60	5472
4	40	4	5	49	1527
5	33	2	2	37	4894
6	10	1	0	11	13374
7	6	1	0	7	6993
8	3	1	0	4	3369
9	17	2	0	19	6108
10	27	6	7	40	1670
Prov.	21	3	2	26	48536

\*"Ceiling values" have been recorded on Laboratory analysis sheets as 600 + and 800 + and are presently being recorded as 999 +. However, on computer cards and tape, they are coded as 601, 801, and 999, respectively.



those samples whose value is known to be above a certain level, but unknown beyond that point.

The prediction of a "feasible" or "probable" value for those samples whose real values are unknown involved derivation of suitable model distributions and the application of simulation techniques. The final prediction depends upon the recorded "ceiling value", the soil area, and the generated random number, in that order. Figure 81 is a diagrammatic flowchart of the simulation of the unknown potassium values. With the aid of this flowchart and Tables XXI, XXII, XXIII, XXIV, and XXV one can follow the method of simulation that was used in the computer program. For example, if a "ceiling value" of 601 occurred for a sample in Area 1, the computer would refer to row 1 of Table XXI. The table shows that for Area 1, 45% of the samples will fall between 601 and 800 lb/ac inclusive; 27% will fall between 801 and 1000 lb/ac, inclusive; and 28% between 1001 and 1600 lb/ac, inclusive. The decision into which interval the sample will fall, will then depend upon the value of the generated random number, which can fall anywhere between 0 and 100, inclusive. If the generated number is between 0 and 45, inclusive, Table XXIV will be referred to and another random number generated to determine the interval within which the final value will fall. Referring back to Table XXI, if the generated random number is between 46 and 72, inclusive, Table XXV will be referred to, and so on. Thus, by following the pattern illustrated by Figure 81 and the related tables, one can perceive a picture of the type of simulation that is taking place. In the long run, the values that are being predicted will form a distribution similar to that of the model used.

Tables XXIII, XXIV, and XXV show the number of samples used to derive



Figure 81. Diagrammatic flowchart of the computer simulation of unknown potassium values.

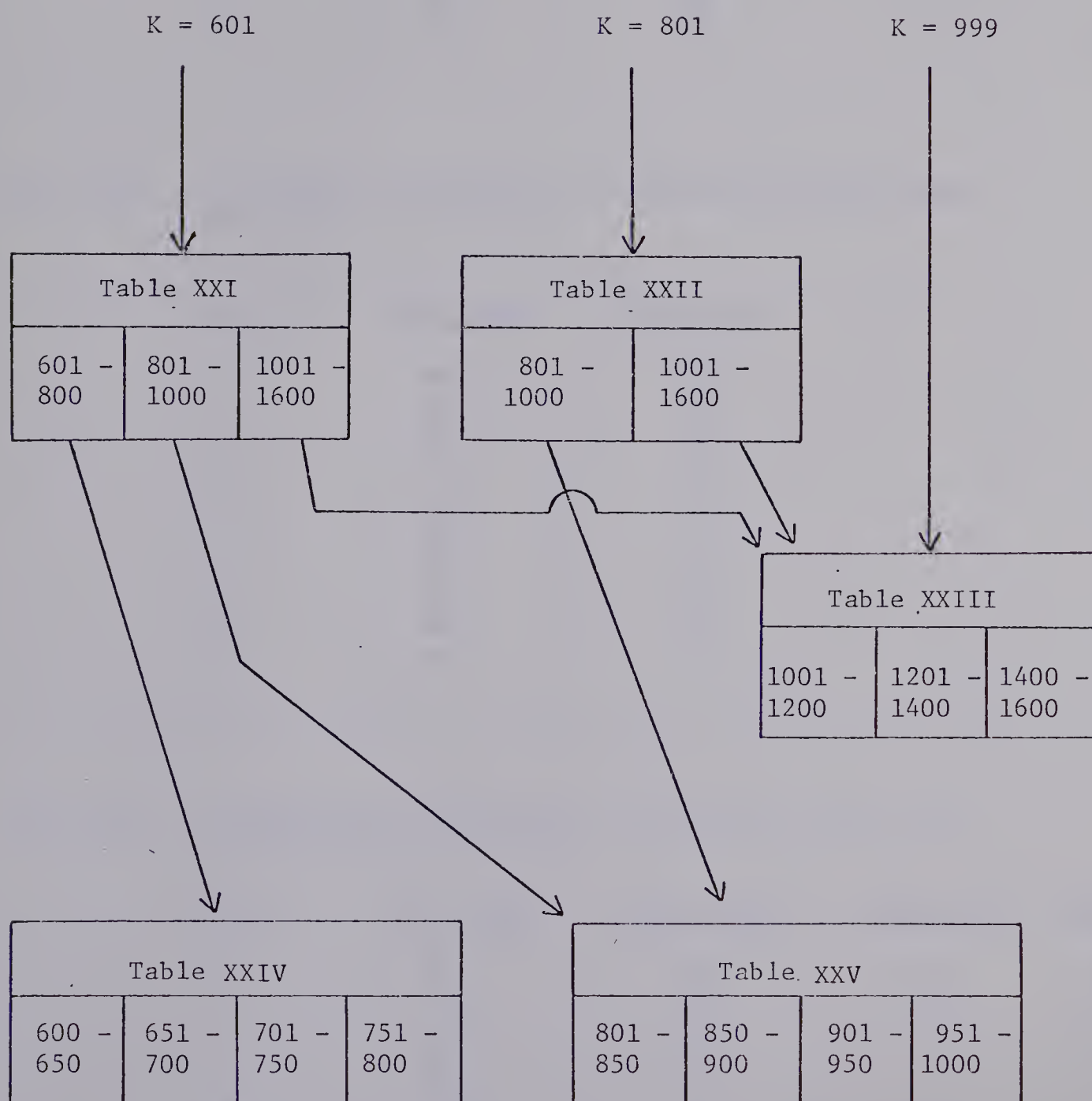




Table XXI. Percentage distribution of potassium values above 600 lb/ac by soil area.

<u>Area</u>	<u>601-800</u>	<u>801-1000</u>	<u>1001-1600</u>
1	45	27	28
2	45	27	28
3	47	23	30
4	57	21	22
5	66	19	15
6	81	15	4
7	84	11	5
8	87	10	3
9	67	18	15
10	49	45	6

Table XXII. Percentage distribution of potassium values above 800 lb/ac by soil area.

<u>Area</u>	<u>801-1000</u>	<u>1001-1600</u>
1	48	52
2	50	50
3	43	57
4	50	50
5	55	45
6	80	20
7	67	33
8	80	20
9	56	44
10	89	11

Table XXIII. Percentage distribution of potassium values above 1000 lb/ac by soil area.

<u>Area</u>	<u>1001-1200</u>	<u>1201-1400</u>	<u>1401-1600</u>	<u>*Number</u>
1	53	28	19	11
2	60	26	14	19
3	47	38	15	23
4	85	15	-	-
5	87	13	-	-
6	100	-	-	-
7	100	-	-	-
8	100	-	-	-
9	83	17	-	-
10	74	26	-	12

\* Number of samples used to derive the distribution





Table XXIV. Percentage distribution of potassium values between 601 and 800 lb/ac, inclusive.

<u>Area</u>	<u>601-650</u>	<u>651-700</u>	<u>701-750</u>	<u>751-800</u>	* <u>Number</u>
1	22	25	25	28	403
2	23	23	27	27	480
3	22	24	25	29	814
4	34	24	22	20	208
5	36	24	22	18	628
6	38	29	20	13	616
7	41	25	20	14	299
8	36	34	15	15	108
9	35	28	22	15	424
10	22	24	29	25	370

\* Number of samples used to derive the distribution

Table XXV. Percentage distribution of potassium values between 801 and 1000 lb/ac, inclusive.

<u>Area</u>	<u>801-850</u>	<u>851-900</u>	<u>901-950</u>	<u>951-1000</u>	* <u>Number</u>
1	28	29	27	16	145
2	41	27	16	16	179
3	28	32	21	19	217
4	36	25	25	14	52
5	44	27	15	14	94
6	35	34	20	11	85
7	40	30	20	10	22
8	40	30	20	10	6
9	50	20	20	10	46
10	37	27	26	10	178

\* Number of samples used to derive the distribution.



the model distributions. The percentage distribution of potassium values over 1000 lb/ac is the least accurately constructed distribution because of the small number of samples used to derive the distribution. It should be noted that these distributions are by no means final. As more results are collected, more precise models can be constructed.

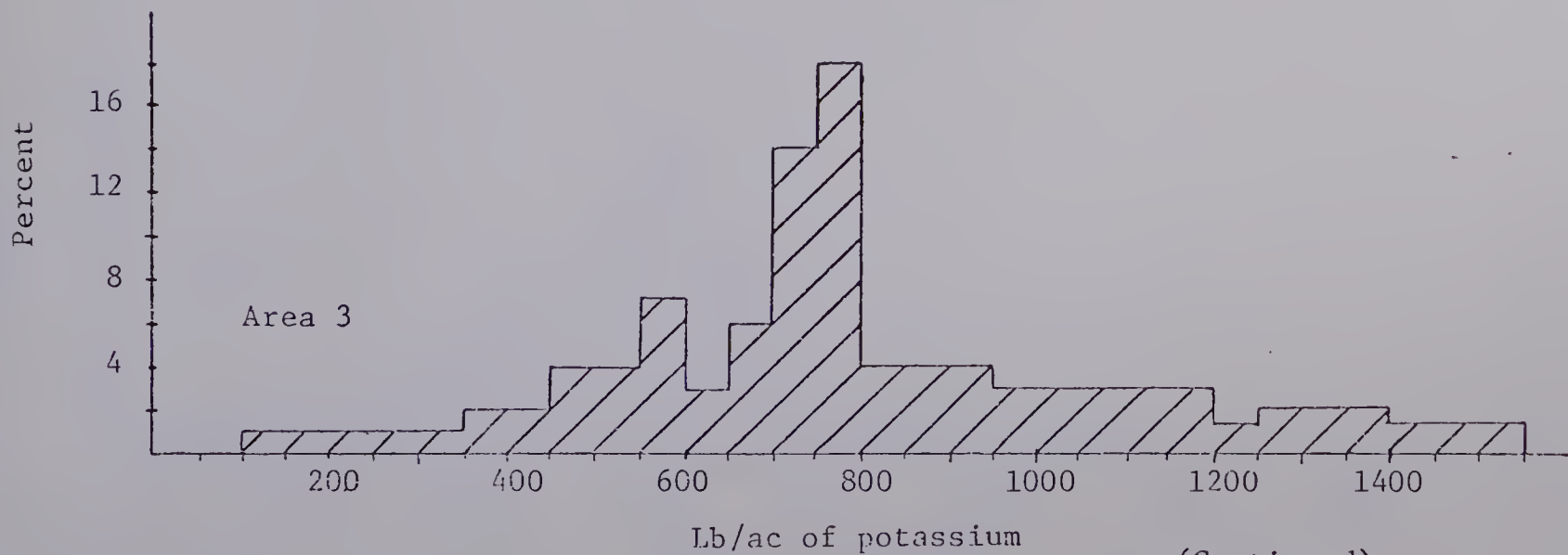
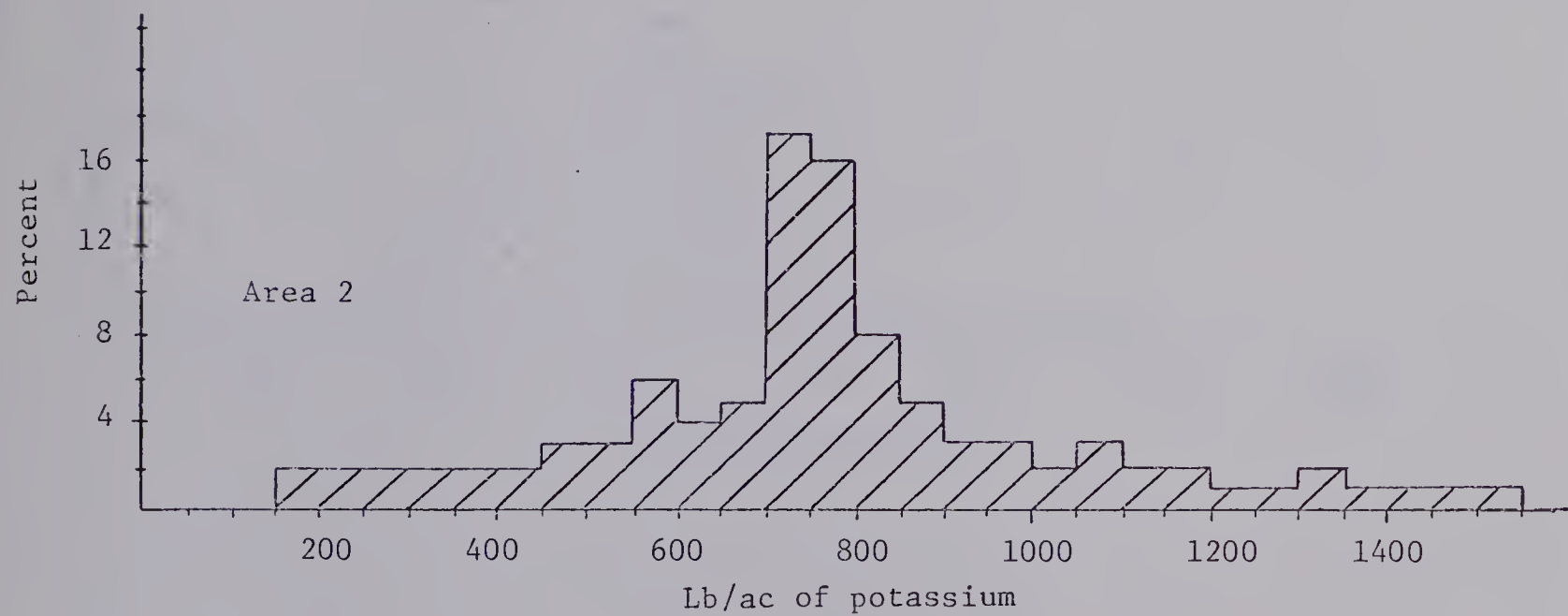
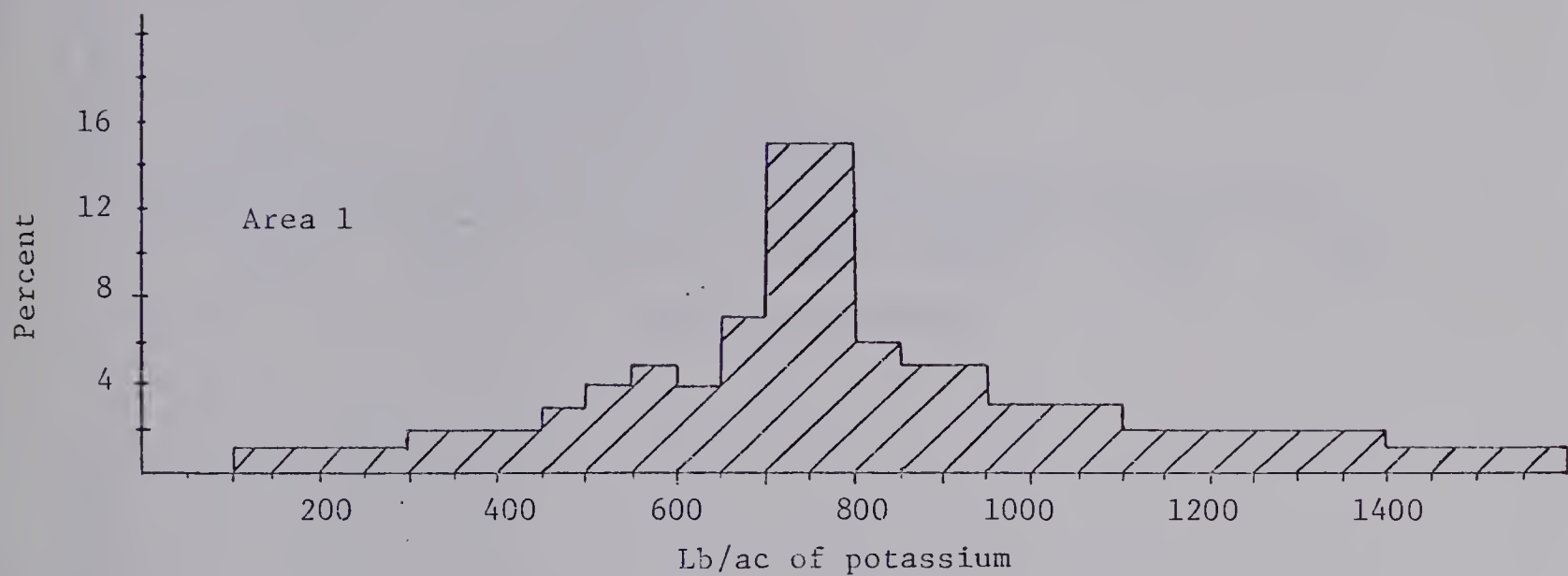
The results in Figure 82 display approximately normal distributions for each soil area. Figure 82 shows the potassium percentage distribution by soil area for all the A.S.F.T.L. samples collected since April, 1965 up to March, 1969 with "ceiling values" replaced by "probable" values as derived from the model distribution. Both the shape and range of the distributions tend to vary from area to area. Areas 1, 2, and 3 show high peaks between 700 and 800 lb/ac. Areas 4, 5, and 6 do not show the same high peaks, but rather, tend to have flatter distributions which are skewed to the left. Both Areas 7 and 8 have distributions skewed to the right. The Peace River area shows the flattest distribution with no outstanding peaks.

Although the values that are simulated are not as accurate as the values that would have been obtained from complete chemical analysis, they are definitely more realistic than the previously recorded "ceiling values".

Generally, the results presented in this section indicate two major facts. Firstly, an accurately predicted value for unknown potassium values is a necessity for impartial analyses of the results. Secondly, as a result of the simulation of "probable" values for those unknown samples, the final distributions appear to be normal and reasonable, and definitely more precise than any distribution formed by using "ceiling values" as real numbers.



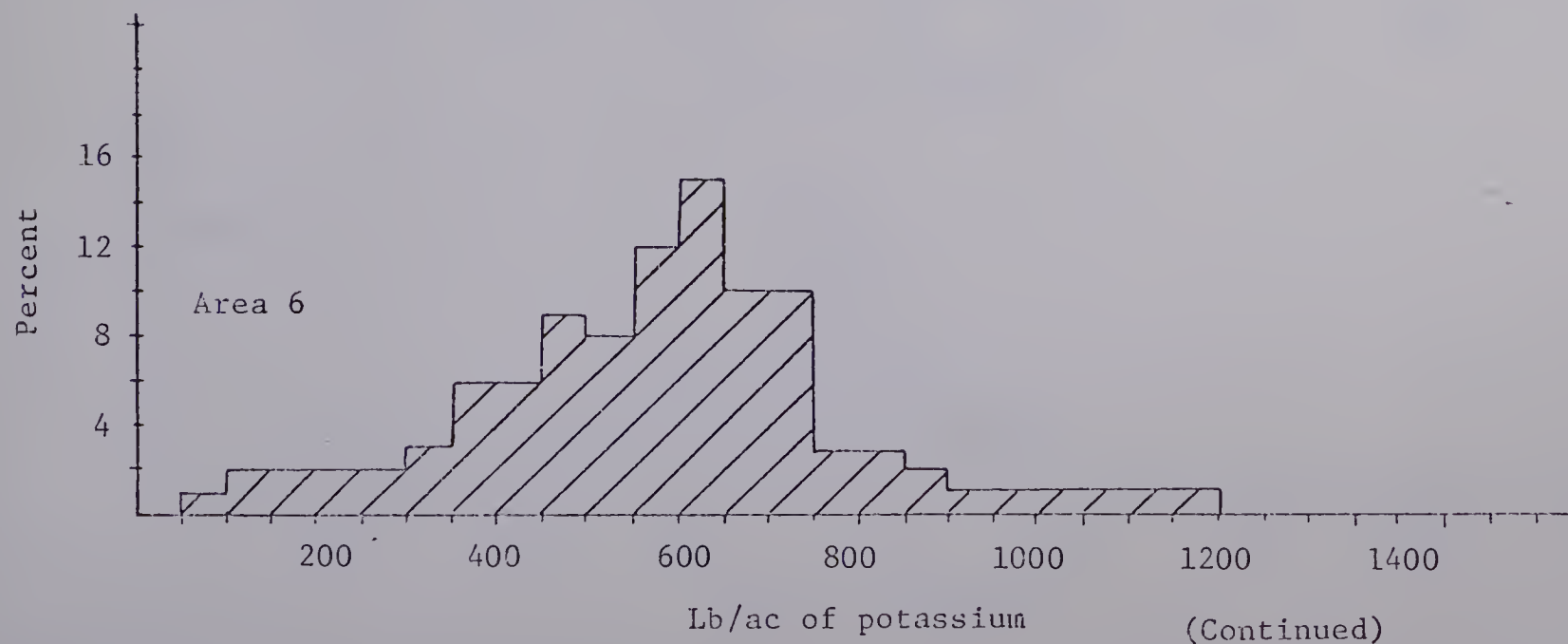
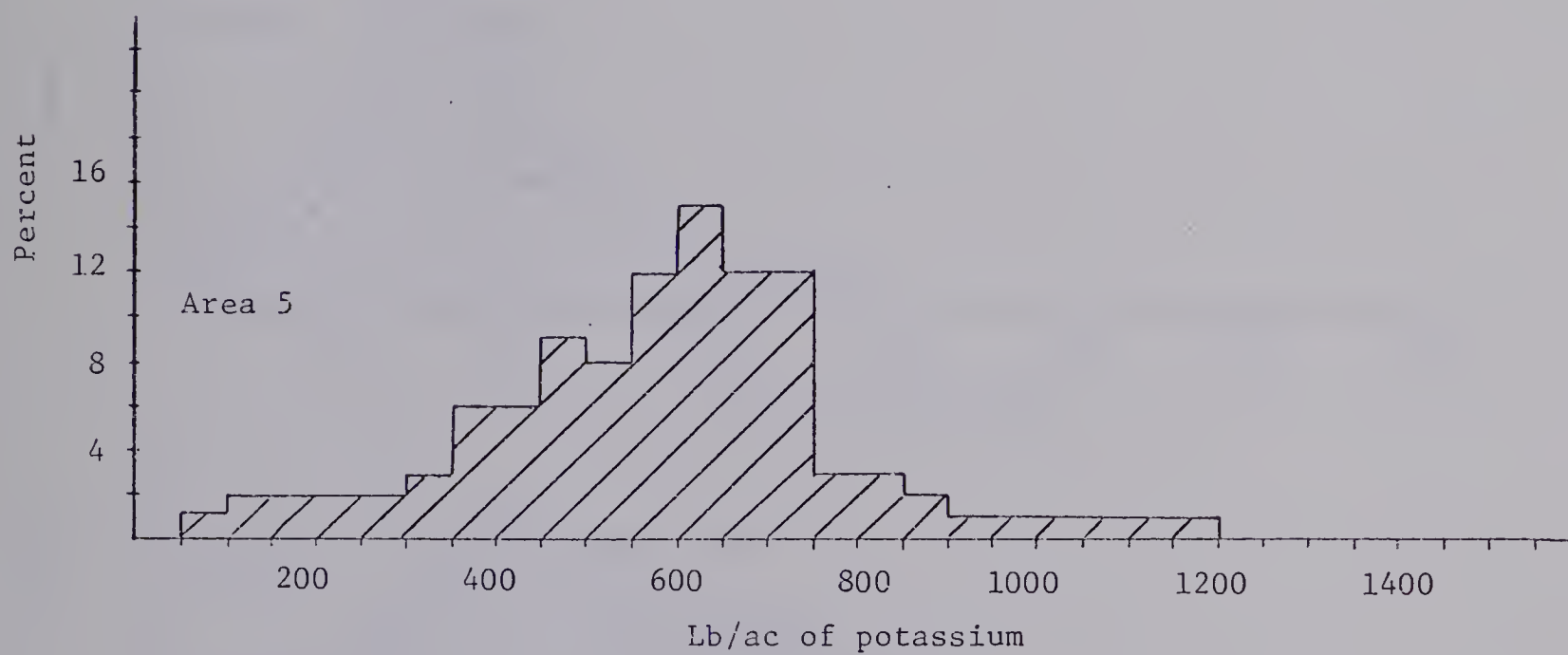
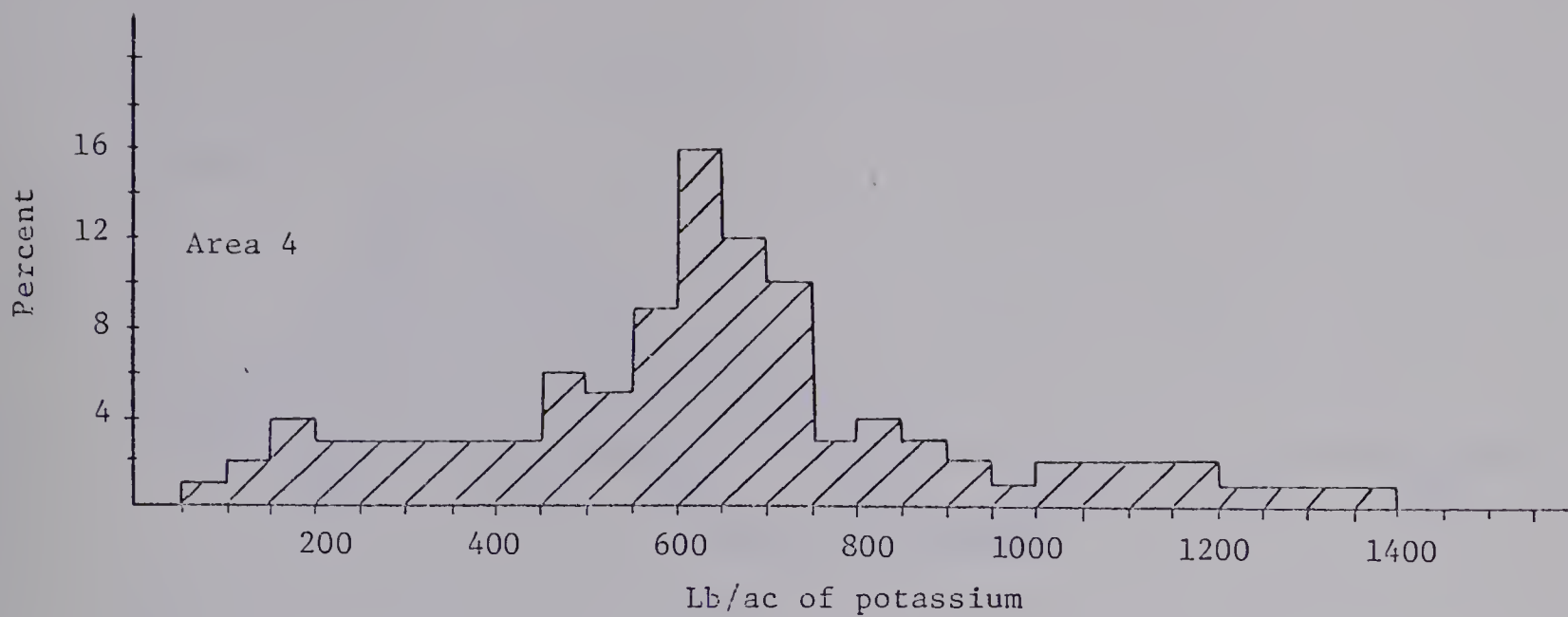
Figure 82. Potassium percentage distribution by soil area.  
Time period: 1965-69.



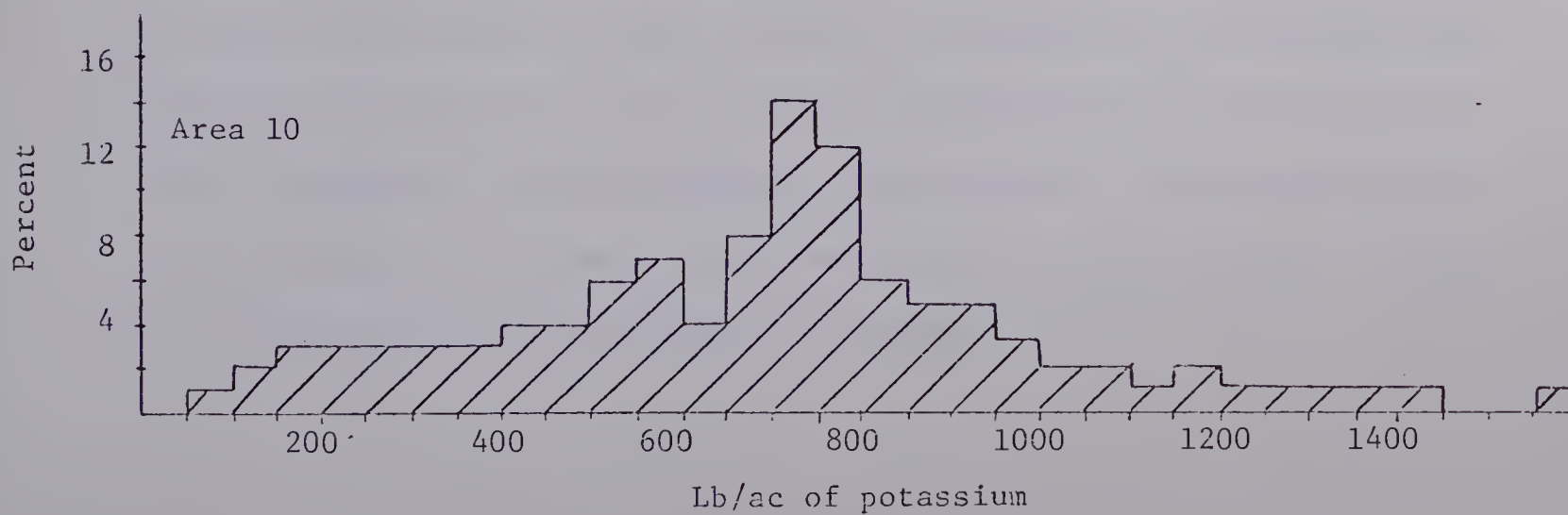
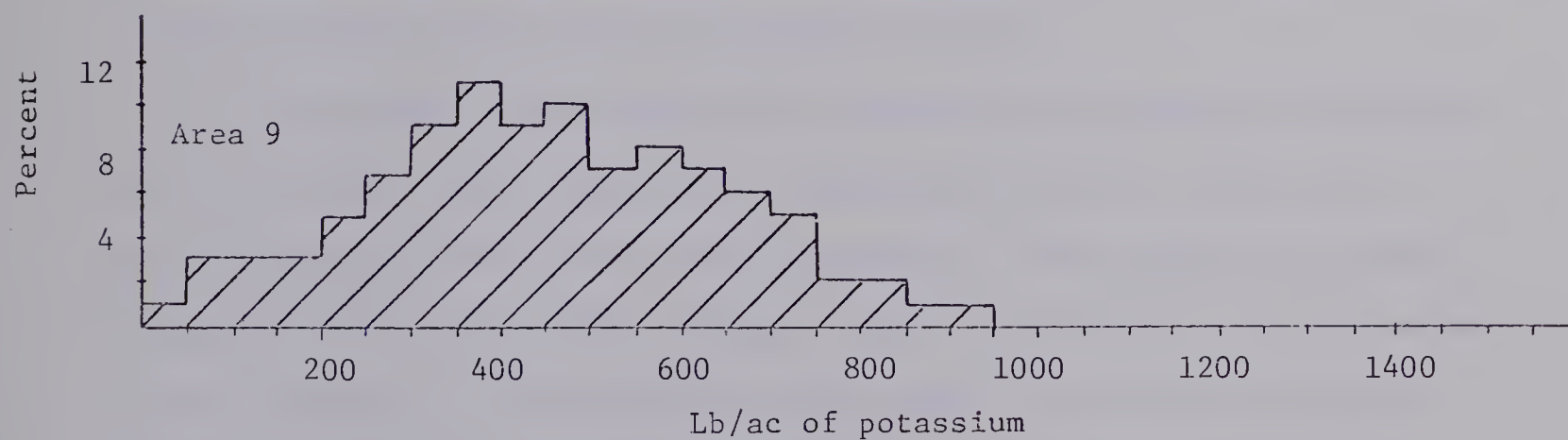
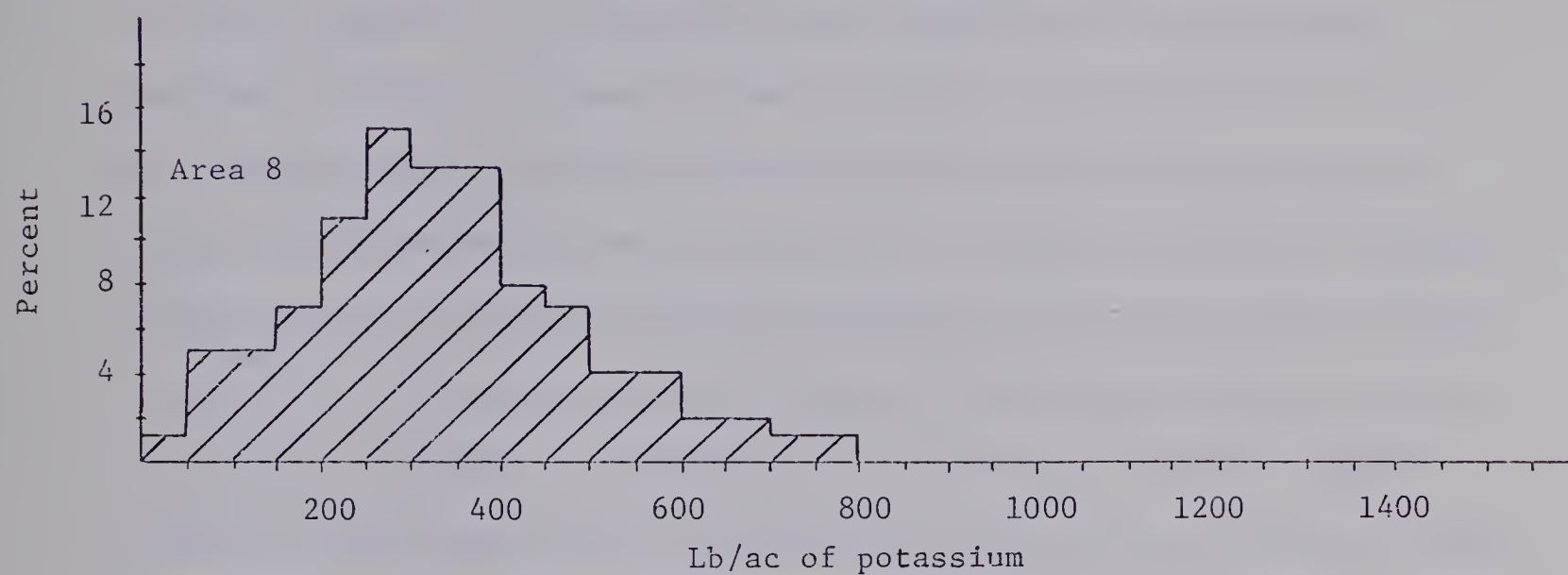
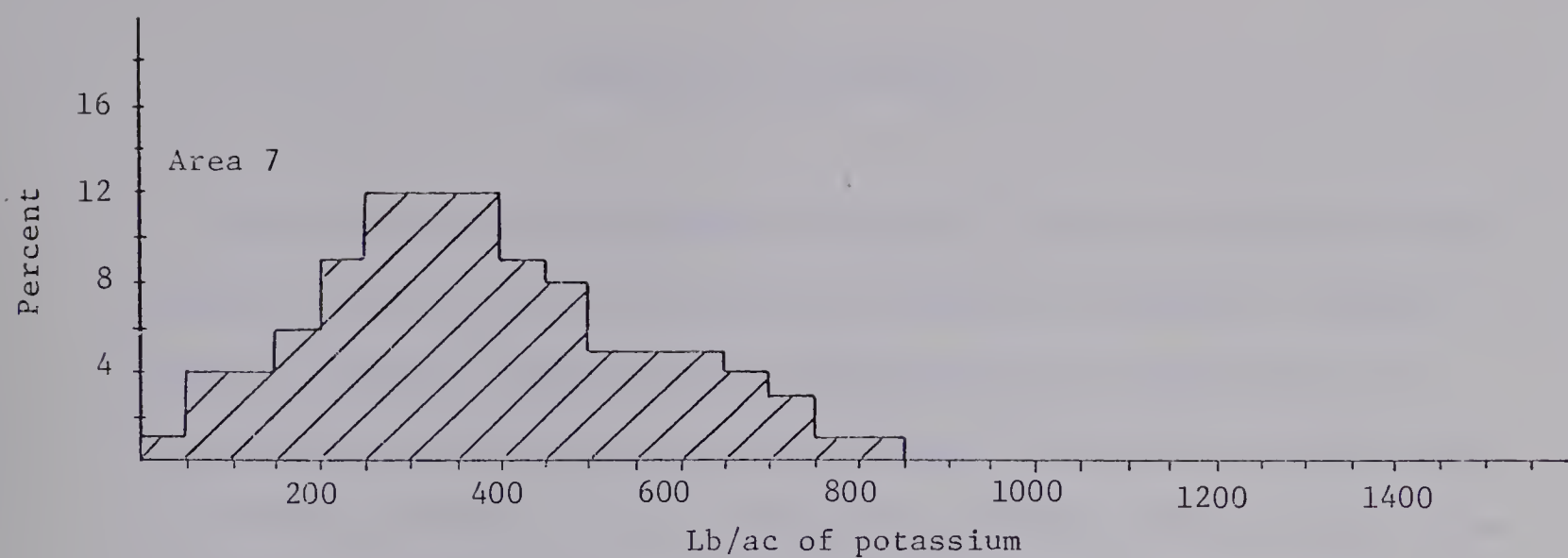
(Continued)













## V. SUMMARY AND CONCLUSIONS

The soil test data available since April, 1962 were transferred from filed records to IBM computer cards and eventually to magnetic computer tape. As of March, 1969, results from approximately 62,500 farm soil samples had been stored on tape. The available information was then subjected to an intensive series of summary-type analyses with the aid of the electronic computer. The summaries were conducted for two principal reasons. Firstly, to present a clear, concise, and accurate picture of the trends existing in the A.S.F.T.L. data for nitrate nitrogen, phosphorus, potassium and pH for the province of Alberta, taking into consideration variation of nutrient levels on different crops and within different time periods, and constructing and making available the computer programs to do this. Secondly, to determine the feasibility of future computer analysis of soil data based on efficiency of the computer programs to manipulate such data and portray the results obtained.

A preliminary investigation was conducted into accuracy of sampling soil test data. The results from a representative field showed that for the provincial average of 16 cores composing a single sample, the sample mean will fall within  $\pm 22\%$ ,  $\pm 20\%$ ,  $\pm 11\%$ ,  $\pm 2\%$  of the true field mean 80% of the time for N, P, K, and pH, respectively. Field size was found to have very little effect on the accuracy of sampling. On a regional basis, size was also not the important factor in determination of accuracy, but rather the number of samples present. The results for field sampling are based on only one field and further work should be carried out to verify the accuracy of the limits referred to above.



Programs to summarize the data for nutrients and pH by crop type and time period for the 10 soil-climatic areas and for the counties and districts in Alberta were developed. The programs proved to be applicable and the results useful, especially when displayed graphically, in depicting general trends present within the data.

Nitrate nitrogen showed considerable variation from year to year, particularly on fallow. Fallow always maintained the highest levels of nitrate with stubble and grass-legume crops displaying lower means. Phosphorus trends were dominately related to soil area, but yearly variations could be noted. The southern areas of Alberta showed the lower phosphorus levels with the northern areas, especially the Gray Wooded soils to the northeast of Edmonton, showing high levels of phosphorus. Soils from most cropping practices appeared to have similar levels of phosphorus, except for vegetable crops where the level was usually exceedingly high. Grasslands tended to show lower levels of phosphorus. Potassium levels were very closely related to the soil-climatic areas and showed little annual variation. Crop type did not appear to significantly affect potassium levels. Soil pH showed trends similar to those of potassium.

As a result of a fairly detailed investigation into the efficiency and applicability of various mapping techniques to the A.S.F.T.L. data, two averaging techniques, the profile method and the circle method, were accepted as the most promising. Of these the circle method was selected because it appeared to be the more accurate mapping function. The circular mapping function was then applied in a computer program to the available soil test data, and plotted nutrient maps of the province were obtained and contoured. The finished maps displayed clearly marked trends. The computer program





developed and utilized to map the province allowed considerable user control of the final outcome, and at the same time, proved to be very efficient and rapid. In less than 4 minutes, using an IBM 360/67 computer, 62,500 records could be scanned, the township totals tallied, the circular mapping function applied to the whole province, and the results transferred to a plot-tape for later plotting.

The trends depicted on the contoured maps followed very closely those shown by the area summaries and county maps prepared but they were, of course, more detailed and accurate than the area and county maps. Nitrate nitrogen on fallow usually showed a higher level in the areas corresponding to the Black soil zone, but this same trend could not be shown for every year. Nitrate levels on stubble and grass-legume crops did not always fluctuate in the same direction as did fallow levels. Stubble levels tended to remain more constant over the years than did fallow levels of nitrate. Phosphorus trends were fairly clear and easy to follow. The area of Solonetzic soil to the southeast of Edmonton could be easily distinguished on the map. Contrary to the nitrate maps, similar trends could be depicted from year to year. These trends, however, did change shape and even relative value from one year to the next. No positive research was carried out to determine the nature of this apparent shifting from year to year.

Relative levels of potassium, when contoured, closely followed the soil colour zones of Alberta, showing a steady decrease from southern Alberta to central Alberta. Some yearly variation did appear, but not any substantial amount. The pH maps showed some variation over the years, but the trends were definitely more closely related to location than annual



variation.

The A.S.F.T.L. data contained "ceiling values" for both nitrogen and potassium which, particularly in the case of potassium, would have caused considerable bias in the results. Therefore, model distributions were devised and "most probable" values were derived by the process of simulation to replace the "ceiling values". The final results proved to be much more accurate than use of the original "ceiling values".

In conclusion, large amounts of valuable soil test information have been summarized and mapped in a very short time with the aid of the electronic computer. This machine has proved to be not only very efficient but essential in the analyses of the vast quantity of soil test data being collected by the A.S.F.T.L. The future applications of the computer to other soil data appear to be unlimited.

The study has suffered from two serious drawbacks. In spite of the large number of samples there were many parts of the province which were very poorly represented and not represented at all. Time will overcome this problem, as more data are accumulated. A second weakness has been the concentration on only the top six inches of soil. The interpretation of all data in this thesis must be done with this in mind. Data for the top two or even three feet would be needed for the most meaningful rating of the nutrient status of Alberta soils.

A worthwhile sequel to this research project would be an attempt to relate the maps of nutrient levels to the soils -- their parent materials, profile types, and classification.



## VI. BIBLIOGRAPHY

- Allmaras, R.R. 1965. Bias. In Methods of soil analysis, part I. Agronomy 9. C.A. Black (ed.). Amer. Soc. Agron., Wis.
- Anderson, R.L. and Bancroft, T.A. 1952. Statistical theory in research. McGraw-Hill Book Co., Inc., N.Y.
- Anderson, R.L. and Houseman, E.E. 1942. Tables of orthogonal polynomial values extended to  $N = 104$ . Iowa Agr. Exp. Sta. Res. Bull. 297, 595-672.
- Bayrock, L.A. and Pawluk, S. 1967. Trace elements in tills of Alberta. Can. J. Earth Sci. 4, 597-607.
- Bell, A.P. and Thornton, S.F. 1937. The effect of season and fertilization on results of rapid chemical tests. Soil Sci. Soc. Amer. Proc. 2, 167-71.
- Bennett, C.A. and Franklin, N.L. 1954. Statistical analysis in chemistry and the chemical industry. John Wiley and Sons, N.Y.
- Black, C.A. 1965. Operator variation. In Methods of soil analysis, part I. Agronomy 9. C.A. Black (ed.). Amer. Soc. Agron., Wis.
- Bowser, W.E. 1967. Agro-climatic areas of Alberta. Alberta Soil Survey. Canada Department of Agriculture.
- Bowser, W.E. et al. 1962. Soil survey of the Edmonton sheet (83-H). Alberta Soil Survey, Report No. 21.
- Box, G.E.P. 1954. The exploration and exploitation of response surfaces: some general considerations and examples. Biometrics 10, 16-60.
- Carolus, R.L. and Lucas, R.E. 1943. Some factors influencing fluctuations in acidity during extreme changes in moisture content of soils. Proc. Amer. Soc. Hort. Sci. 42, 507-10.
- Churchman, C.W. et al. 1957. Introduction to operations research. John Wiley & Sons, N.Y.
- Cline, M.G. 1944. Principles of soil sampling. Soil Sci. 58, 275-278.
- Cochran, W.G. 1953. Sampling techniques. John Wiley & Sons, N.Y.
- Davies, O.L. 1954. Design and analysis of industrial experiments. Hafner Publ. Co., N.Y.





- Dechering, F.J.A. 1956. Objects, aims and importance of soil testing. In The organization and rationalism of soil analysis, part 2. The European Productivity Agency of the Organization for European Economic Co-operation, Paris.
- DeLury, D.B. 1950. Values and integrals of the orthogonal polynomials up to  $n = 26$ . Univ. Toronto Press, Toronto, Can.
- Dixon, J.W. 1965. Extraneous values. In Methods of soil analysis, part I, Agronomy 9. C.A. Black (ed.). Amer. Soc. Agron., Wis.
- Doughty, J.L. 1941. The advantages of soil paste for routine pH determinations. Sci. Agr. 22, 135-138.
- Favret, A.G. 1965. Introduction to the digital computer applications. Reinhold Publishing Corporation, N.Y.
- Ferrari, T.J. and Vermeulen, F.H.B. 1956. Soil heterogeneity and soil testing. In The organization and rationalism of soil analysis, part 2. The European Productivity Agency of the Organization for European Economic Co-operation, Paris.
- Filinger, G.A. 1931. Sampling orchard soils for nitrate determinations. Amer. Soc. Hort. Sci. Proc. 28, 515-518.
- Flagle, C.D. 1960. Simulation techniques. In Operations research and systems engineering. Flagle, C.D. et al. (ed.). John Hopkins Press, Baltimore.
- Goettel, A.W. 1962. The potassium status of some Alberta soils. M.Sc. Thesis, Univ. of Alberta, Edmonton, Alta.
- Grant, F.S. 1954. A theory for the regional correction of potential field data. Geophysics 19, 23-45.
- Griffin, W.R. 1949. Residual gravity in theory and practice. Geophysics 14, 39-56.
- Guenther, W.C. 1965. Concepts of statistical inference. McGraw-Hill Book Co., N.Y.
- Hammond, L.C., Pritchett, W.L. and Chew, V. 1958. Soil sampling in relation to soil heterogeneity. Soil Sci. Soc. Amer. Proc. 22, 548-552.
- Harper, H.J. 1924. The accurate determination of nitrates in soils. Indus. and Engin. Chem. 16, 180-183.
- Hemingway, R.G. 1955. Soil sampling errors and advisory analysis. J. Agr. Sci. 46, 1-8.



- Henry, J.L. 1967. Sampling patterns for a soil test. Report of the meeting of the National Soil Fertility Committee. Saskatoon, Sask. Unpublished.
- Hester, J.B. and Shelton, F.A. 1933. Seasonal variation of pH in field soils - a factor in making lime recommendations. J. Amer. Soc. Agron. 25, 299-300.
- Hoel, P.G. 1966. Elementary statistics. 2nd ed. John Wiley & Sons, Inc., N.Y.
- Holland, D.A., Little, R.C., Allen, M. and Dermott, W. 1967. Soil and leaf sampling in apple orchards. J. Hort. Sci. 42, 403-17.
- Horton, H.A. and Stinson, F.A. 1939. Investigations in sampling soils previously fertilized for flue-cured tobacco. Sci. Agr. 19, 616-621.
- Jacob, W.C. and Klute, A. 1956. Sampling soils for physical and chemical properties. Soil Sci. Soc. Amer. Proc. 20, 170-172.
- Kemphorne, O. and Allmaras, R.R. 1965. Errors of observation. In Methods of soil analysis, part I. Agronomy 9. C.A. Black (ed.). Amer. Soc. Agron., Wis.
- Krumbein, W.C. 1955a. Experimental design in the earth sciences. Trans. Amer. Geophys. Union 36, 1-11.
- Krumbein, W.C. 1955b. Statistical analysis of facies maps. J. Geol. 63, 452-70.
- Krumbein, W.C. 1956. Region and local components in facies maps. Bull. Amer. Assoc. Petr. Geol. 40, 2163-2194.
- Krumbein, W.C. 1959. Trend surface and contour-type maps with irregular control-point spacing. J. Geophys. Res. 64, 823-834.
- Krumbein, W.C. and Miller, R.L. 1953. Design of experiments for statistical analysis of geological data. Jour. Geology 61, 510-532.
- Leo, M.W.M. 1963. Heterogeneity of soil sampling. J. Agr. Food Chem. 11, 432-434.
- Mader, D.L. 1963. Soil variability - a serious problem in soil-site studies in the northeast. Soil Sci. Soc. Amer. Proc. 27, 707-709.
- McCollum, R.E. and Nelson, W.L. 1954. How accurate is a summary of soil testing information. Soil Sci. Soc. Amer. Proc. 18, 287-292.
- Melsted, S.W. 1967. The philosophy of soil testing. In Soil testing and plant analysis, part I. Soil Sci. Soc. Amer., Wis.



- Nelson, R.W. 1964. Alberta soils with low potassium values. M. Sc. Thesis, Univ. of Alberta, Edmonton, Alta.
- Newton, J.D. 1954. Microbial maintenance of nitrogen in Western Canada's grey-wooded, black earth and brown prairie soils. Transactions of the 5th International Congress of Soil Science, Leopoldville, 3, 76-87.
- Newton, J.D., Wyatt, F.A. and Brown, A.L. 1945. Effects of cultivation and cropping on the chemical composition of some Western Canada Prairie Province soils. Part 3. Sci. Agr. 25, 718-737.
- Odynsky, W.M. 1962. Soil zones of Alberta as established by Alberta soil surveys. 3rd ed. Research Council of Alberta.
- Oldham, C.H.G. and Sutherland, D.B. 1955. Orthogonal polynomials: their use in estimating the regional effect. Geophys. 20, 295-306.
- Parker, F.W., Nelson, W.L., Winters, E. and Miles, J.E. 1951. The broad interpretation and application of soil test information. Agron. J. 43, 105-112.
- Peck, T.R. and Melsted, S.W. 1967. Field sampling for soil testing. In Soil testing and plant analysis, part I. Soil Sci. Soc. Amer., Wis.
- Peevy, W.J., Smith, F.V. and Brown, P.E. 1940. Effects of rotational and manurial treatments for twenty-one years on the organic matter, nitrogen, and phosphorus contents of Clarion and Webster soils. J. Amer. Soc. Agron. 32, 739-753.
- Petersen, R.G. and Calvin, L.D. 1965. Sampling. In Methods of soil analysis, part I. C.A. Black (ed.). Agronomy 9. Amer. Soc. Agron., Wis.
- Prince, A.L. 1945. Determination of total nitrogen, ammonia, nitrates and nitrites in soils. Soil Sci. 59, 47-52.
- Protz, R., Presant, E.W. and Arnold, R.W. 1968. Establishment of the modal profile and measurement of variability within a soil landform unit. Can. J. Soil Sci. 48, 7-19.
- Raupach, M. 1951a. Studies in the variation of soil reaction II. Seasonal variations at Barooga, N.S.W. Austr. J. Agr. Res. 2, 73-82.
- Raupach, M. 1951b. Studies in the variation of soil reaction III. Variations at Waite Agr. Res. Inst. Austr. J. Agr. Res. 2, 83-91.
- Reed, J.F. and Rigney, J.A. 1947. Soil sampling from fields of uniform and non-uniform appearance and soil types. J. Amer. Soc. Agron. 39, 26-40.





- Rice, W. and Riehm, H. 1956. Sampling. In The organization and rationalism of soil analysis, part 1. The European Productivity Agency of the Organization for European Economic Co-operation, Paris.
- Riehm, H. 1956. Methods of soil sampling and preparation of samples. In The organization and rationalism of soil analysis, part 2. The European Productivity Agency of the Organization for European Economic Co-operation, Paris.
- Rigney, J.A. 1956. Sampling for composition studies. Amer. Soc. Hort. Sci. Proc. 68, 569-574.
- Rigney, J.A. and Reed, J.E. 1946. Some factors affecting the accuracy of soil sampling. Soil Sci. Soc. Amer. Proc. 10, 257-259.
- Rouse, R.D. and Bertramson, B.R. 1949. Potassium availability in several Indiana soils: its nature and methods of evaluation. Soil Sci. Soc. Amer. Proc. 14, 113-123.
- Sampford, M.R. 1962. An introduction to sampling theory with applications to agriculture. Oliver and Boyd, Edinburgh, London.
- Shickluna, J.C. and Walsh, L.M. 1967. Application of electronic data processing equipment to soil testing. In Soil testing and plant analysis, part I. Soil Sci. Soc. Amer., Wis.
- Shiue, Cherng-Jiann and Chin, N.L. 1957. Direct use of pH values in statistical analysis of soil reactions. Soil Sci. 84, 219-224.
- Snedecor, G.W. 1956. Statistical methods. The Collegiate Press Inc., Ames, Iowa.
- Steel, R.G.D. and Torrie, J.H. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., N.Y.
- Stein, C. 1945. A two-sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Stat. 16, 243-258.
- Sukhatme, P.V. 1954. Sampling theory of surveys with applications. Iowa State College Press, Ames, Iowa.
- Swanback, T.R. and Morgan, M.F. 1930. Seasonal variations in soil reaction. Conn. Agr. Exp. Sta. Bull. 311, 264-268.
- Thomas, S. 1965. Computers. Holt, Rinehard and Winston, Inc., N.Y.
- Tisdale, S.L. and Nelson, W.L. 1966. Soil fertility and fertilizers. 2nd ed. The Macmillan Co., N.Y.





- Toogood, J.A. and Anderson, G.H. 1964. Report of Alberta Soil Test Appraisal Project. ARDA Project No. 9016.
- Toogood, J.A., Bentley, C.F., Webster, G.R. and Moore, A.W. 1962. Grey wooded soils and their management. Dept. of Soil Sci. Univ. of Alta. Bull. S-M-I.
- Van Bavel, C.H.M. 1967. Use and abuse of information processing by machine. In The collection and processing of field data. A CSIRO Symposium. E.F. Bradley and O.T. Denmead (ed.) Interscience Publishers, N.Y.
- Vermeulen, F.H.B. 1953. Control of the reproducibility and accuracy of routine analyses at the laboratory for soil and crop testing in the Netherlands. Plant and Soil 4, 267-275.
- Walker, P.H., Hall, G.F. and Protz, R. 1968. Soil trends and variability across selected landscapes in Iowa. Soil Sci. Soc. Amer. Proc. 32, 97-101.
- Walsh, L.M. 1964. Computers tailormade recommendations. Crops Soils 17 (1), 15-17.
- Welch, C.D. and Fitts, J.W. 1956. Some factors affecting soil sampling. Soil Sci. Soc. Amer. Proc. 20, 54-56.
- Whitten, E.H.T. 1959. Composition trends in granite:modal variation and ghost stratigraphy in part of the Donegal granite, Eire. J. Geophys. Res. 64, 823-834.
- Wyatt, F.A., Ward, A.S. and Newton, J.D. 1926. Nitrate production under field conditions in soils of Central Alberta (1925-26). Sci. Agr. 7, 377-385.
- Wyatt, F.A., Ward, A.S. and Newton, J.D. 1926. Nitrate production under field conditions in soils of Central Alberta. Sci. Agr. 7, 1-24.



## VII. APPENDICES

## APPENDIX I. Format of A.S.F.T.L. Data

Card 1

<u>Columns</u>	<u>Information</u>
1-2	Year
3-4	Month
5-6	Day
	} Date received
7-11	Report number
12-16	Blank
17-18	Lowest sample number
19	Card number
20	Blank
21-50	Surname and initials
51-80	Address (box and street number)

Card 2

<u>Columns</u>	<u>Information</u>
1-18	Same as Card 1
19	Card number
20	Blank
21-50	Town or city
51-78	District Agriculturist office
79-80	Province



Card 3

<u>Columns</u>	<u>Information</u>	
1-11	Same as Card 1	
12-16	Laboratory number	
17-18	Real sample number	
19	Card number	
20	Blank	
21-48	Name of crop to be grown	
49-50	Blank	
51-52	Crop code    table number	
53-54	Percent legume	
55	Irrigation	
56-57	Number of places sampled	
58-60	Field size (acres)	
61	Erosion	
62	Fall fertilizer	
63-64	Drilled fertilizer	} Present fertilizer history
65-67	Pounds per acre	
68-69	Broadcast fertilizer	
70-72	Pounds per acre	
73-79	Blank	
80	Call back number 1	





Card 4

<u>Columns</u>	<u>Information</u>
1-18	Same as Card 3
19	Card number
20	Blank
21-22	Previous crop code number
23-24	Percent legume
25-27	Yield of crop
28	Units of yield
29-30	Drilled fertilizer
31-33	Pounds per acre
34-35	Broadcast fertilizer
36-38	Pounds per acre
39	Damage
40	Cause
41	Management
42	Seed down
43-40	Blank
51-52	Quarter
53-54	Section
55-57	Township
58-59	Range
60	Meridian
61-62	Type of district
63-65	Number of district

Past fertilizer  
history

Legal location

(Continued)



<u>Columns</u>	<u>Information</u>
66-67	Year
68-69	Month
70-71	Day
72-79	Blank
80	Call back number 2

Card 5

<u>Columns</u>	<u>Information</u>
1-18	Same as number 3
19	Card number
20	L
21-23	Nitrogen (lb/ac)
24-26	Phosphorus (lb/ac)
27-29	Potassium (lb/ac)
30	Sodium
31-32	Soil pH
33-35	Conductivity
36	Sulphate
37-38	Area
39	Organic matter
40	Lime
41	Texture

Top 6 inches  
of soil

(Continued)

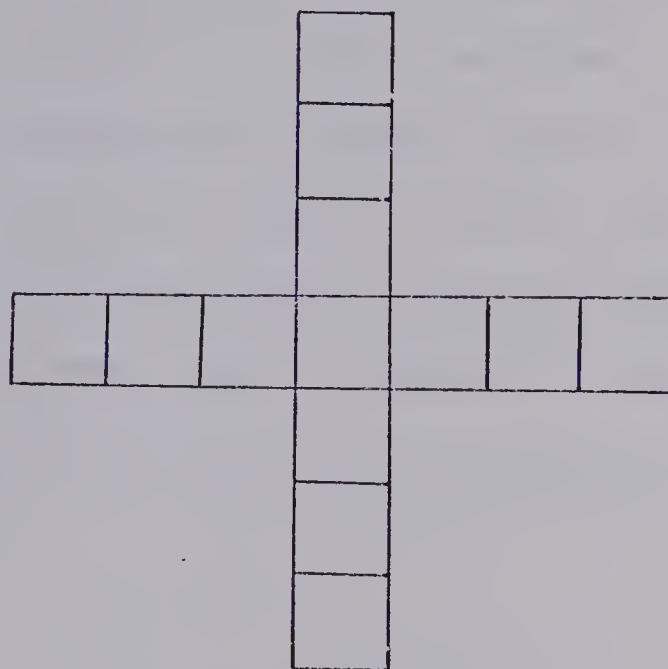


<u>Columns</u>	<u>Information</u>	
42-44	Nitrogen (lb/ac)	6-12 inches of soil
45-47	Phosphorus (lb/ac)	
48-50	Potassium (lb/ac)	
51	Sodium	
52-53	Soil pH	
54-56	Conductivity	
57	Sulphate	
58	Organic matter	
59	Lime	12-18 inches of soil
60	Texture	
61-63	Nitrogen (lb/ac)	
64-66	Phosphorus (lb/ac)	
67-69	Potassium (lb/ac)	
70	Sodium	
71-72	Soil pH	
73-75	Conductivity	
76	Sulphate	
77	Organic matter	
78	Lime	
79	Texture	
80	Call back number 3	



APPENDIX II a. Diagram and description of the profile  
mapping technique

The profile mapping technique was essentially a cross-shaped averaging function which can be diagrammatically illustrated as shown below.



Each individual square represents a township or grid point. The central township is the one for which a prediction or map value calculation is being made. The computer program takes into account all the townships within each of the four arrays of the mapping function, including the central township. An average of values within these townships is used to form a final map value. Weights can be assigned to the various townships, where the central township and those townships closest to the central township are weighted heaviest. If this is done, the final calculation would be a weighted average. Since the function is usually being moved over a





large number of townships, the name weighted moving average is often more descriptive.

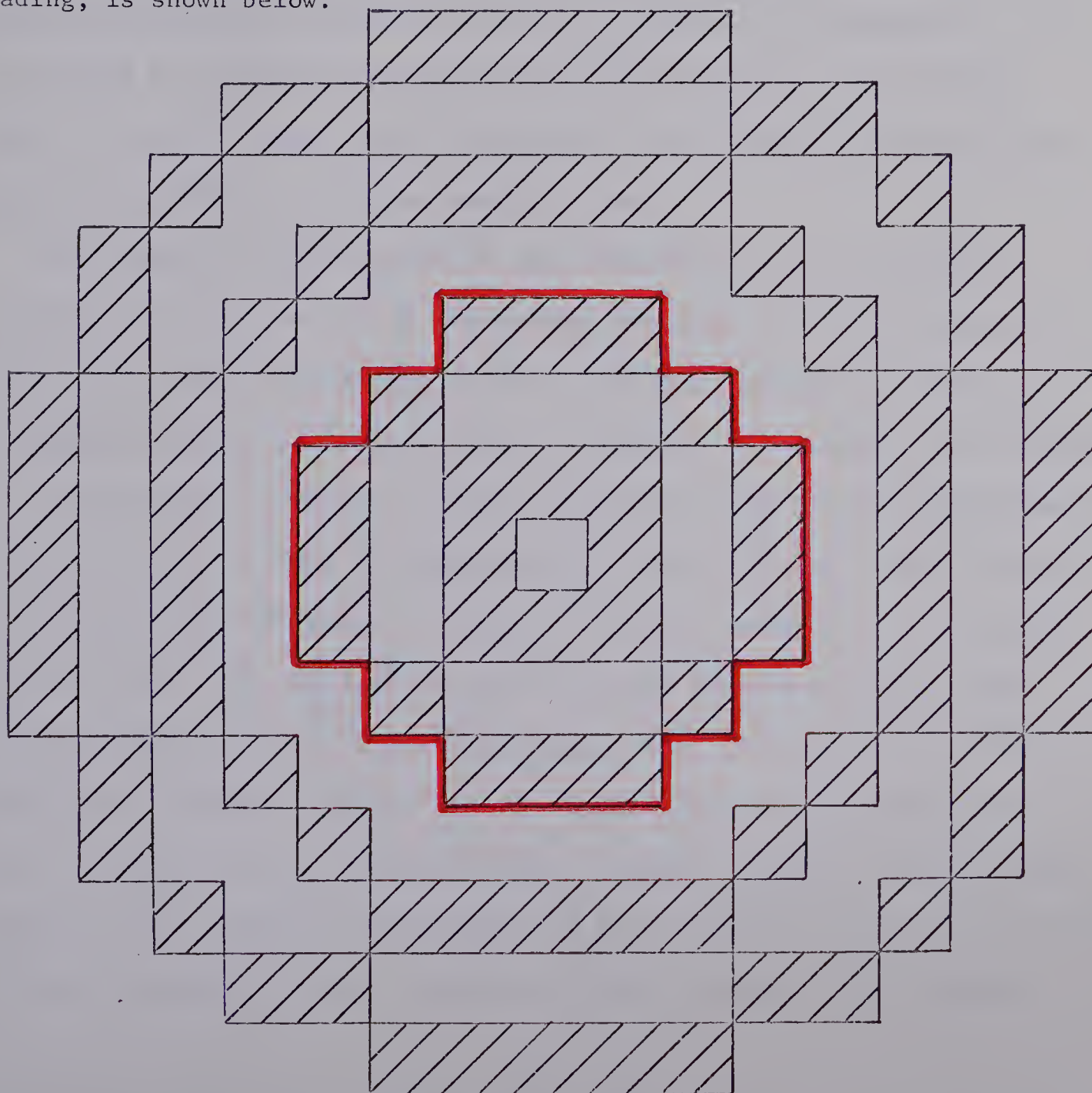
The technique can be used when average values per township are present, when totals per township are present, or when weighted values for each township are present. The discussion in Appendices IIb and III should clarify what is meant by these 3 categories of use.

The length of each of the arrays as well as the assigned weights can be varied depending upon the degree of smoothness required. In addition, the computer program permits a critical distance in the x- and y- directions to be set up. If the number of samples within this critical distance is less than the particular number of samples considered necessary to make an accurate calculation, then no calculation is made. The number of samples considered necessary rests on the judgement of the user.



APPENDIX II b. Diagram and description of the circular mapping technique.

The circle method is based on the same averaging technique as the profile method except that rather than a straight line of townships in the east-west and north-south direction, the circle method is amplified to take into account concentric rings of townships about the grid point to be mapped. The description "pseudo-concentric" rings would probably be more appropriate since townships will not form true circles. A diagram of the circle function used, with its' 8 concentric circles portrayed by shading, is shown below.





The method allows the size of the circle used for averaging to be controlled by elimination of the concentric circles not wanted. Any weights desired in order to depict the regional trend may be assigned to the circles. The sample numbers in the individual townships within each concentric circle will act as weights and weight the final result accordingly, unless otherwise specified. If an average value per township is used, all townships with samples available will be equally weighted.

A critical radius may be assigned arbitrarily and will include a certain number of inner circles. The critical radius or the critical circle that has been used for the majority of the maps presented in this manuscript contains 4 concentric circles (a total of 37 townships). It is outlined in red in the diagram on the preceding page. An expulsion number or critical number may be assigned to the critical circle and predictions not made where too few samples exist.

The computer program works on each concentric circle separately starting from the inner circle and working outward to the 8th concentric circle. The center circle lies on the township or grid point that is to be mapped and therefore includes only 1 township. The second circle includes the 8 surrounding townships, the third circle the next 12 surrounding townships and so on. A value is calculated for each concentric circle outward to the last concentric circle or until the weight assigned to a concentric circle is zero. The critical circle is usually smaller than the actual mapping circle, although it may be the same size. It is therefore usually reached first and the total number of samples (or some representation of this) is compared to the critical number required. If less, no prediction is made for the township being mapped and the averaging function moves to an adjacent township and begins the whole procedure again. If the number





required is present, the calculation continues until the value for the last concentric circle has been calculated. These values are then averaged using the weights assigned to the concentric circles as the weighting factors. Thus a weighted average is obtained for each township where the critical number of samples are available. If the townships are weighted according to the number of samples they contain before the mapping function is called, the final mapped value will be the result of two separate types of weighting. The actual mapping function used for mapping the soil nutrient status of Alberta soils has three levels of weighting which can be used. The technique is called moving because it is moving from one grid point to another as it makes map value calculations.



### APPENDIX III. Description of the Alberta mapping program.

The final product of the Alberta mapping program is a plotted map of Alberta showing areas of high and low values of the particular nutrient being mapped. Each map required, on the average, less than 4 minutes execution time and approximately 24 minutes of plotter time using the IBM 360/67 and the Calcomp plotter available in the Computing Center, Department of Computing Science, University of Alberta.

The computer program begins by reading the data cards containing the control variables and then the magnetic tape containing the A.S.F.T.L. data. A check is made at each record scanned for proper years, months, and crops according to the control variables on the data cards. Those samples which do not have the proper qualifications of crop or time of sampling are passed by.

Individual samples that are accepted can be weighted according to the number of places sampled in the field. This particular weighting is based on the fact that the ability of a sample to adequately represent a field depends primarily on the number of individual cores composing the sample. As the number of randomly sampled cores increases, so does the accuracy. In addition, the program allows the elimination of those samples based on too few cores. However, if none of this is desired, the sample weighting and elimination schemes can be by-passed, and every sample included as it is.

The program then continues converting township and range values into x- and y- grid coordinates. These coordinates are used to store the township totals as well as the number of samples received per township. At



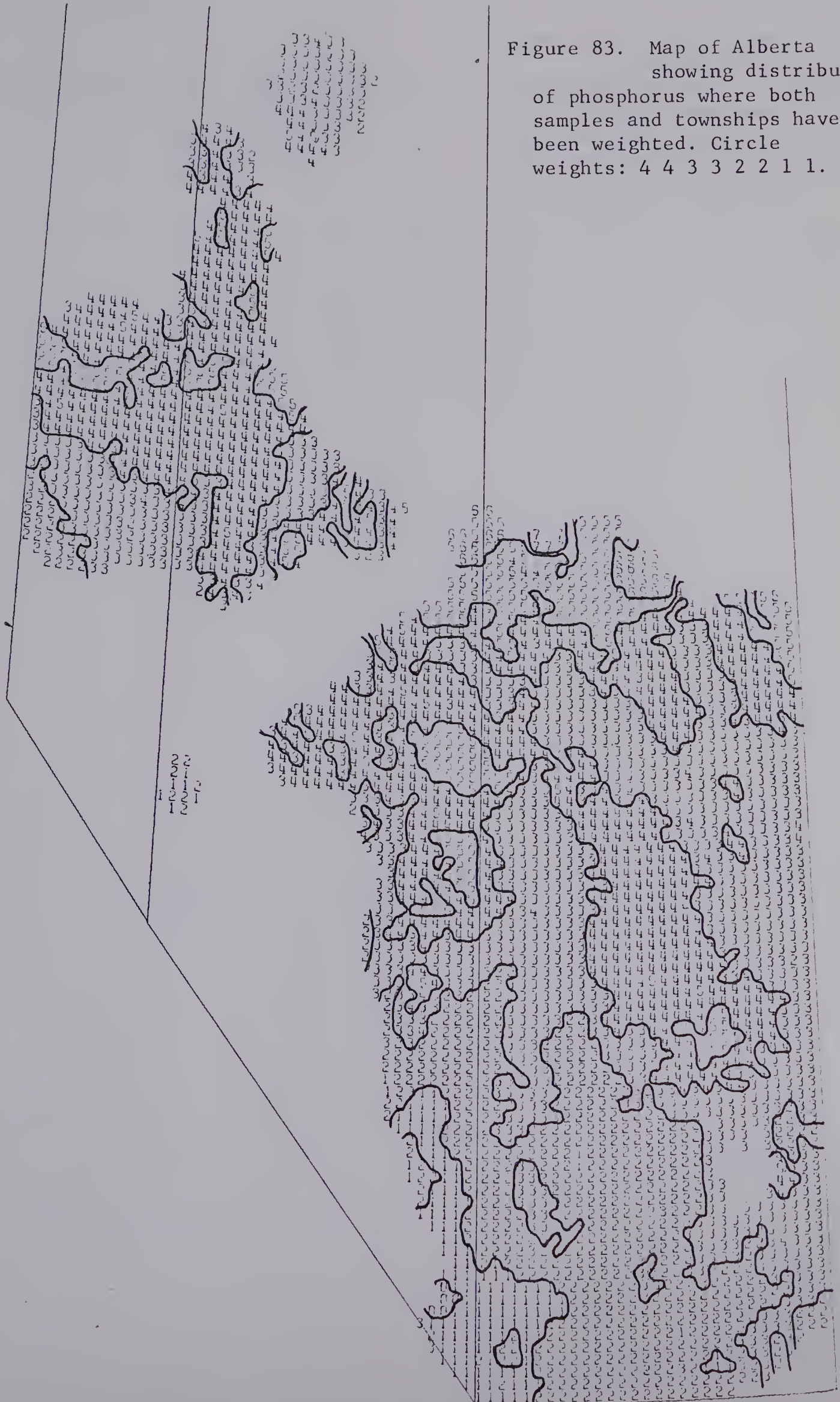
the same time they enable rapid movement of the circular mapping function over the province.

When all the records required have been read and the values stored in their appropriate locations, then, depending upon the control variables previously read in, a weighting may be set for each township based on the number of samples within each township, or an average can be obtained for each township having samples (thus, automatically setting the weighting to one for each township with samples), or township totals and numbers may be left as they are (thus automatically letting the number of samples per township act as the weighting for each township). The program variables also allow the elimination of those townships with very few samples available.

The weightings or no weightings or eliminations on both the field level and township level, allow many combinations of control. However, the results show thus far, that very little difference can be noted for maps with different field and township weighting controls, but within the same time period and with the same crops considered. For example, Figures 83 and 84 are maps of Alberta showing the distribution of phosphorus for the same time period (1962-1969) and on all crops, but with different field and township control variables. Figure 83 shows the full extent of control. Samples are assigned a weighting based on the number of cores composing them, and similarly townships are weighted based on the number of samples within each township. Figure 84, on the other hand, shows the effect of little control where the control variables assign no weights and eliminate no samples. Not even the number of samples within each township act as weights, because an average for each township is calculated

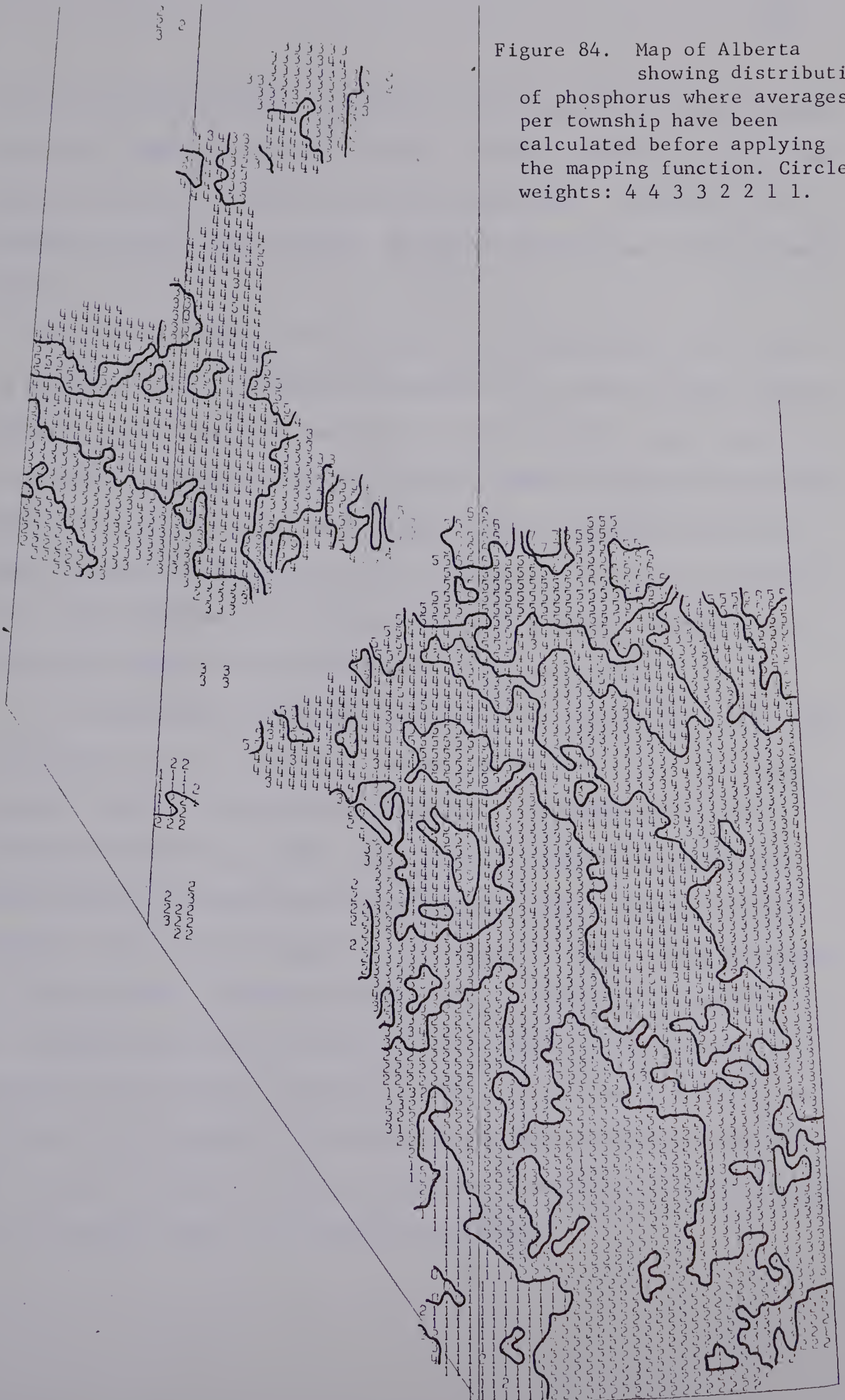














before the mapping function is applied. The two maps show almost identical trends. Where differences do exist, the more accurate map would most likely be the one in Figure 83 with full weightings. Obviously, if unreasonable weights were assigned, then quite different maps would probably result.

The next major step in the program is the application of the circular mapping function, discussed in Appendix IIb, to the grid-like representation of the province formed by the legal locations. Map values are predicted for townships where the required number of samples is available within the critical circle. The weights that are assigned to the 8 concentric circles were previously read in as part of the control variables. These weights and the size of the circle are the primary factors in controlling the degree of smoothness of the final map surface.

The area between the 4th and 5th meridian is mapped first. The townships next to the 4th meridian are automatically lined up by their legal location. However, those townships along the 5th meridian are not automatically lined up by their legal location values. In the map of Alberta, partial townships exist along the east side of the 5th and 6th meridian. The computer program is constructed so that if these townships are less than half a township in size, they automatically become part of the township directly to the east of them. Because of this gradual disappearance of townships, "correction lines" must be made to readjust the location of the townships. These are not the same correction lines that occur every 24 miles, but rather, major corrections that are necessary for accurate mapping about the 5th and 6th meridians.



Therefore, after completion of mapping between the 4th and 5th meridians, the townships are shifted in such a way as to align those townships that are next to the 5th meridian. As stated above, the alignment of the townships along the 5th meridian into a straight line of column grid values is necessary because of the disappearance of townships. Without this alignment, calculation of the mapped values would be biased. A similar type of shifting for mapping accuracy is performed about the 6th meridian.

The final maps produced have one other form of bias which is no fault of the program, but rather the available data. All the maps produced suffer from what is known as the "edge effect". This is particularly true next to the borders of the province where no samples are available in half of the mapping circle, thus heavily slanting the final value to one side. A similar effect can probably be found for predictions in the fringe areas.

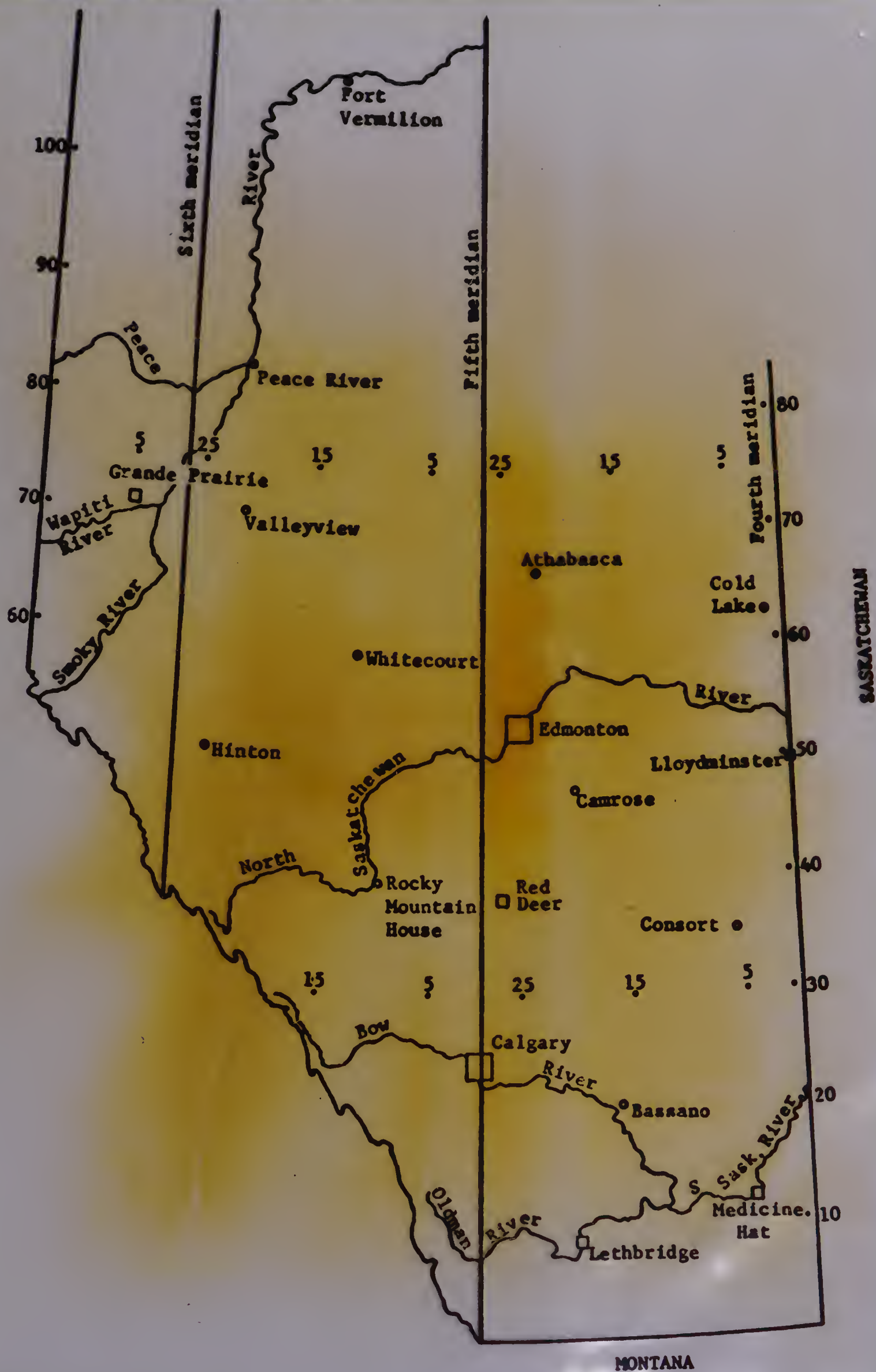
After the mapping function has completed the mapping of the province, the values can be rounded-off into intervals for easier presentation of the results. The results are then transferred to a plot-tape for later plotting of the soil nutrient map developed. The final plotted map is a polyconic projection centered about the 5th meridian (114°). The plot can be of any size, not exceeding the size of the plotter paper.















**B29906**